Title
Essays On The Economics of Agriculture, Information, and Climate Change

Author
Trilnick, Itai

Publication Date
2019

Peer reviewed|Thesis/dissertation

Essays On The Economics of Agriculture, Information, and Climate Change

by

Itai Trilnick

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David Zilberman, Chair
Professor Michael Anderson
Professor David Anthoff

Summer 2019

Essays On The Economics of Agriculture, Information, and Climate Change

Abstract

Essays On The Economics of Agriculture, Information, and Climate Change

by

Itai Trilnick

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor David Zilberman, Chair

Agricultural practices are increasingly dependent on precise, real-time weather information. Irrigation, pest management, and even strategies for climate change adaptation require adequate information inputs. This dissertation starts with a valuation of weather information in California, assessing some of the the economic gains from the California Irrigation Management Information System (CIMIS). This system of weather stations and information portal was intended for water saving in irrigation, but turned out to have many unexpected uses. The next chapter uses CIMIS historical information to estimate the yield response of California pistachios to warm winters. This has been an ongoing challenge in the literature, as the available yield data are scarce and noisy. Merging them with CIMIS data is essential in estimating this response, and for predicting the yield effects of winter temperatures in the future. The last builds on these predictions, and analyzes the potential gains from a technology that could deal with the challenge of warming winter days: spraying the dormant trees with kaolin clay, which reflects part of the sunlight and keeps the trees cooler. To assess the gains, I take the estimated response function, climate predictions, and other variables to market model that is solved numerically and allows for welfare gain calculations. Beyond the potential of this specific technology for pistachios, tweaking temperature distribution tails seems like a promising concept for climate change adaptation. Many temperature challenges brought by climate change are in these distribution tails, where yield responses are often non-linear. This means that small temperature adjustments, local in time and place, might suffice for cost-effective adaptation strategies. I call this concept "Micro-Climate Engineering", and predict that such practices will be increasingly popular with the progression of climate change.

To my grandparents

I started the PhD program at UC Berkeley in 2013 with three living grandparents. I am graduating in 2019 with none. Abuela Rosa passed away when I was very young, and Abuelo Raúl made it until my last year here. Jaime ("Zeide") and Clara ("Bobe") departed while I was navigating through classes and research. None of them got to see my graduation ceremony, but I know they were happy and proud that I was doing my PhD at Berkeley. *Abuelas y abuelos, gracias por su amor y cariño ¡Los quiero mucho!*

Contents

List of Figures

List of Tables

Acknowledgments

This dissertation is the final product of a six year journey. I am lucky to have had the chance to study at a great department in one of the world's best universities. My advisor David Zilberman was very insightful and encouraging. His unique views of economics and life in general were always refreshing. He also helped me leverage a project with the California Department of Water Resources to a full dissertation (and a few years of rent and tuition). I want to thank my dissertation committee and the rest of the faculty at Cal who spent time teaching me, listening to my ideas, and adding comments and suggestions to my work. I also want to thank Carmen Karahalios and Diana Lazo from the department of Agricultural and Resource Economics for logistical (and emotional) support, going the extra mile and beyond that for the students. I have had the privilege of meeting, studying, and exchanging ideas with brilliant peers who have become good friends. I am very happy that Yotam Shem-Tov, my friend from Israel, decided to start his program at the Economics department the same time I joined ARE. Besides being roommates for a year, we spent a lot of time together studying, discussing research, or just chatting and relaxing. I am glad to have witnessed him and Juliana Londoño-Velez become a couple. John Loesser was one of my first local friends, and helped me understand some quirks not only in econometrics but in American culture as well. Ed Rubin and Hilary Soldati graduated early, proving that one can be not only an economist and a nice person, but also have a family and still be more efficient than I am. Deirdre Sutula, Tamma Carleton, Erin Kelly, Greg Lane, Matt Woreman, Elizabeth Ramirez, Danamona Andrianarimanana, the rest of my cohort and other students at ARE were all inspiring, each in their own way. Keeping sane on the PhD track was partly thanks to the Cal Karate Club, let by Sensei Hidetaka Abe. On a broader perspective, this PhD is the latest achievement I owe to the love and care of my parents since I was born. I thank them and my brother for listening to me and always giving good advice. Last but not least, my greatest accomplishment at Berkeley is getting Angel Ryono to say "yes". She has taught me so much, and her loving support helped push through in difficult moments. I look forward to spending whole weekends with her once I graduate.

Chapter 1

Introduction

1.1 The value of weather information in agriculture

Weather is a key input for agricultural production. A vast economic literature is dedicated to the role of weather information in grower decision making, market outcomes, and commodity futures. On one hand, better information (and forecasts) about the weather can help growers optimize their use of other inputs, increasing efficiency in production and avoiding costs related with uncertainty. On the other hand, some economic models can show—under some assumptions—that more precise weather information might not be welfare increasing, as *ex-ante* uncertainty about the weather can lead to extra investment in other inputs. That is, when growers have better forecast of adverse weather, output would be further reduced from its level under uncertainty (Lave, 1963). There is also some concern about weather forecasts acting as signals for collusion among growers, but simple price mechanisms can technically reduce output and welfare with better weather prediction even in a competitive market (Babcock, 1990). Notwithstanding these warnings by economists, the economic gains from weather information are usually deemed positive, even if their magnitude is sometimes contested (MacAuley, 2005).

Much of the seminal economic literature on the value of weather information was written between the 1960's and the 1990's, when significant improvements in forecasting was achieved with the advance of computing power and complex meteorology models (Tribbia, 1997). This literature is based on the agricultural practices and available data of that time. While literature about the value of weather information seems to have plateaued in the 2000's, perhaps as forecasting technologies matured and stabilized, the surge of precision agriculture could re-ignite interest in this topic. Heterogeneity within fields and precise growing strategies, based on exact measurement of weather variables (e.g. evapo-transpiration or degree hours), is increasingly the subject of research and technological application (Gordon et al., 2018). Uncertainty regarding real-time weather on micro scales poses conceptually similar questions to those dealt with by the weather forecast literature in the past. At the same time, new discussions on the value of weather information and the government's role in providing it have been revived with advances in remote sensing and satellite technology (Cirac-Claveras, 2019).

The technical and scientific capabilities required to gather and analyze weather data, as well as the non-rival nature of weather information as a product, meant that much of the development of weather services has been done by governments. Johnson and Holt (1997) point out that this led to a significant economic literature, assessing the potential gains from better weather information given the public expenditures. Their survey of the relevant literature mostly includes econometric studies, where the output gains from improved forecasting are estimated and the economic gains from providing them are then calculated per hectare. Other methodologies include survey based valuation, paired with economic data and modeling. Anaman and Lellyett (1996) assess the gains from a weather information system for cotton growers in Australia, finding the benefit-cost ratio of the system at 12.6 (for cotton alone). Klockow, McPherson, and Sutter (2010) conduct a survey based study of the value of the Mesonet network in Oklahoma. Less than 4% of Oklahoma's cropland is irrigated, and the modest value they find for Mesonet information mostly comes from risk management. Interestingly, there are few such examples of an economic study about a specific weather information system in the published literature, as opposed to numerous studies on the value of information for growers. Johnson and Holt do mention, for example, that weather forecast services in Sweden and New Zealand have gone through "extensive privatization", but do not cite any articles analyzing these decisions.

The first part of this dissertation is an analysis of economic gains from the California Irrigation Management Information System (CIMIS), a network of weather stations and data center run by the California Department of Water Resources. For over 30 years, this system has been used by growers, consultants, and other users in California agriculture. This chapter presents the preliminary findings from a thorough report on the value of CIMIS, showing substantial gains not only in agriculture but also in landscape management, regulation, research, and industry.

1.2 Weather information and climate change

Climate change poses a major challenge for agriculture, as predicted shifts in temperature and precipitation patterns around the world affect agricultural productivity (Zilberman et al., 2004; Carleton and Hsiang, 2016). Early studies on climate change in agriculture first focused on the impacts of changing mean temperatures, and more recent empirical literature emphasizes the importance of temperature variance and extreme heat on yields, especially during the growing season (Auffhammer and Schlenker, 2014). For example, Schlenker and Roberts (2009) show sharp drops in the yields of corn, soybean, and cotton, when exposed to degree days above 28–30^0C. Similar findings have been replicated in various crops and locations around the world. Climate scientists affirm that heat waves will increase in frequency and duration as the process of climate change advances (IPCC, 2013). Researching yield responses to high temperatures, especially when the relationship seems non-linear or threshold like, is therefore essential for prediction of climate change effects on agriculture. This can only be done with adequate weather information.

Chapter 3 presents an analysis of the yield response of pistachios to hot winters. This is also a temperature distribution tail problem, at least when looking at temperatures between November and March. Daytime temperatures in California winters have been rising in the past 20 years, and are predicted to rise further in the future. This can have detrimental implications for pistachios, a major California crop, but estimating the yield response function has been a challenge so far. I use CIMIS data and innovative techniques to recover this relationship and predict the potential threat of climate change to California pistachios. It turns out that Pistachios, a billion dollar crop in California, could be threatened by warming winter within the next 20 years.

1.3 Micro-Climate Engineering

While the scope and magnitude of our current climate crisis might be unprecedented in human history, this is not the first time that humans are facing climatic challenges in agriculture. Olmstead and Rhode (2011) show how, through the 19^{th} and 20^{th} centuries in North America, wheat growers managed "...to push wheat cultivation repeatedly into environments once thought too arid, too variable, and too harsh to farm". The transition was made possible mostly by the development of new varieties. Plant breeding toward that end required information on the climate both in the progenitor native areas and the areas where the eventual new varieties would be planted (see Kingsbury, 2009, for a history of plant breeding and its role in agriculture.).

Adaptation to climate can be on the physical dimension as well. Specific interventions can be designed to change the physical environment surrounding plants. The most obvious intervention is building irrigation systems, to compensate for lack of adequate rainfall and soil moisture. But examples of adaptation to temperature by physical means exist as well. This type of intervention is common for a left tail effect: frost. A short lasting fall or spring frost lasts a few hours and can cause substantial damages. To avoid it, only a slight increase in temperature is required, and growers know how to do that.

Some examples for dealing with frost are hundreds of years old. The Tiwanaku civilization formed a system of raised fields on the shores of lake Titikaka in the 7–12 centuries C.E. Fields in select locations were raised with extra soil, up to a few feet above the ground level. Water from nearby springs was diverted and run through canals dug in these raised fields. This provided not only moisture for the plants, but also converted the top soil level into a large heat storage unit. On frost nights, which are common in this high area, the heat stored in the soil kept the near-surface temperatures on raised fields higher than the normal air temperatures, preventing plants from freezing (Kolata and Ortloff, 1989). Without modern weather instruments, the Tiwanaku realized that slight differences in ambient temperatures can have crucial consequences, and planned their fields according to their understanding of the climate. This system yielded far better than regular dry farming practiced before in this area, and supported a larger population than the one residing on the lake shores in the 1990's. Eventually, as climate became drier, the water level of lake Titikaka dropped and the springs dried up, resulting in the collapse of the Tiwanaku culture (Binford et al., 1997). Despite its eventual failure, this technology was successful in abating frost damage for centuries, maintaining a population of hundreds of thousands and showing the power of human intervention on the field level to tackle a temperature distribution tail challenge.

In Europe, traditional methods of dealing with frosts in vineyards include lighting small fires or "frost candles". A more modern approach uses big fans, circulating the cold air in the inverted layer with the warmer air on top of it. Farmers have been using "air disturbance technology" in the US since the 1950's (Hu et al., 2018). Wind generators are used around the world to protect wine grapes, fruits, and even tea from spring frosts. In some cases, a similar effect can be achieved with sprinklers (Olen, Wu, and Langpap, 2015; Lu et al., 2018).

Interestingly, little economic literature has focused on air disturbance technologies. Stewart, Katz, and Murphy (1984) assess the value of weather information in the Yakima Valley of central Washington, in the context of frost prediction and air disturbance technologies. This descriptive study was published in the Bulletin of the American Meteorological Society. Searching the *EconLit* database for "frost" in article titles returns only four results involving actual frost in agriculture, none dealing with temperature altering. A search in the abstracts of papers published by the American Journal of Agricultural Economics results in two papers, neither mentioning air disturbance technologies. The seeming dis-interest in these technologies is even more peculiar in 2019, when weather information is more accessible than ever: air disturbance systems are now sold with online communication to weather services, with the option for automatic operation in case of frost, and can often be switched on and off remotely. They are probably more efficient and valuable than ever before, given advances in technology and the high value of certain frost-sensitive crops.

Technologies such as air disturbance are examples of a concept I call "Micro-Climate Engineering" (MCE). These are relatively small interventions in temperature distributions, limited in space and time, which aim to avoid the nonlinear effects of the extremes. The frost examples discussed above deal with left tail effects. There are also technologies available to deal with right tail effects, which is the focus of my last chapter.

The final chapter of this dissertation deals with an MCE proposal for California pistachios. Chapter 3 deals with the threat of warm winters on pistachios, estimating the potential losses to this high value crop from climate change. Chapter 4 deals with a proposed solution. The MCE technology proposed for this challenge is spraying the dormant trees with kaolin clay, a non-toxic white substance which reflects the sunlight. Sprayed trees have been shown to experience lower temperatures than control trees, and their yields were higher. This intervention requires precise hourly measures of temperature, so growers can track the buildup of special temperature metrics and decide if and how much treatment is required.

Using the pistachio yield-temperature response, estimated in the previous chapter, I build a model that integrates MCE in the pistachio market. The model can be solved with and without the option to use MCE, under various weather realizations. The value of MCE for

California pistachios is calculated as the difference in welfare measures attained in each case. The expected net present value of MCE in pistachios for 2020-2040 is assessed in billions of US dollars. This is yet another example of the potential use of weather information for dealing with climate change challenges in agriculture.

Micro-Climate Engineering might remind some readers of Geo-Engineering, a controversial climate change adaptation concept. Geo-engineering proposals involve global scale interventions in the atmosphere and hydrosphere that would revert some of the changes in the total temperature distribution worldwide (Irvine et al., 2016). In contrast, MCE is a small scale concept, aiming to tweak the temperature tail distributions where necessary rather than shifting the entire distribution year round. Many MCE technologies already exist and are used by growers, making sense both on the technical and economic dimensions. I believe many more examples are out there to be found, and many more will evolve as growers adapt to climate change.

Chapter 2

Assessing The Gains From Public Weather Information: the case of CIMIS

2.1 CIMIS and California agriculture

This chapter assesses the gains from a weather service provided by the California Department of Water Resources (DWR): the California Irrigation Management Information System (CIMIS). Established in 1982, it now comprises of hundreds of weather stations, scattered in many of the growing regions in California, and centralized computing systems for distributing the information and interpolating data between the stations. The intended purpose of CIMIS was to provide accurate real-time information for growers to optimize irrigation and save water. Specifically, many CIMIS stations include evapo-transpiration (ET) sensors, applied on specially maintained turf. Agronomists have been publishing crop coefficients, which serve to transform the turf-based ET measures for use in various commercial crops. This way, growers can estimate how much water has been used by their plants, and plan replenishment of soil moisture accordingly. CIMIS also reports other weather variables, such as temperature, relative humidity, wind speed and direction, and soil temperature at the station. It does not offer forecast services.

CIMIS has become a staple of agricultural practice and research in California. Searching for it on Google Scholar results in 2,860 entries for articles and publications. The gains from CIMIS have previously been analyzed by a team of researchers from the University of California – Berkeley (Parker et al., 1996), and the findings were published widely (Cohen-Vogel et al., 1998; Parker et al., 2000). This report used a survey methodology, and found a 13% applied water reduction with CIMIS, 8% yield increase, and a total economic gain of $32.4 million yearly. The 1996 report also found some examples of unintended use of CIMIS, which in fact delivered a substantial portion of the gains. For example, while the system was mainly designed for improving irrigation performance and water saving, the researching team found that there are many gains from CIMIS use in pest management. CIMIS detailed temperature data are used to calculate pesticide application timing, reducing the amounts of pesticide and increasing yields.

This chapter presents and analyzes the main findings from a more recent report, prepared for DWR by David Zilberman, Itai Trilnick, and Ben Gordon. This report was meant to update the knowledge on the current uses and users of CIMIS, its economic gains, and potential future improvements. The full report is yet in the writing process. However, several patterns and conclusions can already be drawn, and are presented in this chapter.

2.2 Methodologies

The study is based on a survey of CIMIS users. Before running the survey, extensive interviews were held with various users to gather narratives about the roles of CIMIS in different contexts. These interviews provided a first qualitative picture of current CIMIS uses and interactions with other technologies and practices. They suggest that CIMIS has indeed become a mainstay in California agriculture, especially for growers relying on drip irrigation. However, many farmers access CIMIS indirectly through consultants, and might not be aware of the uses and benefits of the system. With the advance of alternative decision making tools (weather stations, soil monitors, etc), CIMIS is now part of a larger information eco-system. The interviews showed that the public availability of CIMIS data, including historic records, are highly regarded among users. This historical and cross-sectional information store is extremely valuable for decision-making and research. It is essential for calibrating other weather tools, verifying their results, and designing water management schemes that require knowledge of the historical distribution of weather variables. In addition, it may even be used to more accurately value farmland.

Interviews were followed by a small survey, carried out by phone and aimed at assessing the initial insights from the interview. The final step was a full scale online survey, sent to all registered CIMIS users. Results from this survey are the basis for economic value calculations.

2.3 Survey details

The electronic survey was designed together with the CIMIS team, considering the results from the initial phone survey. It was decided to try and survey all registered CIMIS users, sending invitations to the email addresses used for enrollment. This might not cover all existing users: some might still be getting information through a third party, such as a consultant or media sources. There are also some electronic services which do not require registration, such as a File Transfer Protocol (FTP) enabled server run by CIMIS. To survey potential users who are not registered, as well as potential users who are not currently using

CIMIS, an invitation to participate was also sent by email to mailing lists provided by the CIMIS team.

The survey included some general questions, directed to all audiences, and questions tailored to specific user groups identified in the initial survey: growers, consultants, users in landscape management, regulators, researchers, and others. The CIMIS team decided that the survey will not include direct questions about water use, costs, and willingness to pay for CIMIS services, especially when addressing growers. These questions were deemed too intrusive, jeopardizing both the response rate and general trust of users in the CIMIS system. This meant that a direct WTP approach, like the ones used in the previous study of CIMIS and the one used by Anaman and Lellyett (1996), would not be possible. Most questions were framed either in Likert-like scales or as a relative response (e.g. percent water saved). The analysis of the results uses indirect assessing of CIMIS impacts, using these types of responses and outside information.

The online survey was done on a commercial platform, Survey Gizmo. It is worth noting that most registered users are not active. In fact, CIMIS user statistics show that a relative small percent of registered users had logged in and extracted data from the system in the year before the survey. Altogether, we have 3,057 responses, out of which 2,358 are complete.

2.4 General Results

The breakdown of self-reported user types is listed in Table 2.1. About 1/4 of our respondents report their primary activity, as it relates to CIMIS, to be agriculture. The second largest category is "other", encompassing a mix of respondents who, in our opinion, should have picked another definition, and a few others who seem to use the data for personal research. This category has gardeners, nursery workers, water consultants, government workers, and a few retired people working on individual projects. We did not reclassify obvious mis-responses, as that would not change the fact that they ended up answering a different set of questions than their "real" category. The third category is government workers, followed closely by research, environmental consulting and landscape management.

About 60% of respondents are aged 45 and above, and only about 17% are aged 25-35. While this might be the result of the age distribution in the major fields of occupation which are potential CIMIS users, it could also be that the current interface of CIMIS caters less to younger potential users who might seek the data elsewhere. About a quarter of respondents are women, and their share decreases at higher age groups. This probably reflects the changing labor force characteristics in CIMIS related professions over the past few decades.

In terms of geographic location, most respondents (80%) report only one area of activity, with the San Joaquin Valley leading the count. Figure 2.1 below shows the shares of respondents in each region. Note that we allowed more than one response for location.

We ask all respondents to rank each type of data, offered by CIMIS, according to the frequency they search for it. Figure 2.2 shows the breakdown of answers for each of the frequency choices. ET and precipitation are large shares of the "often" column. These

User Type	Count	Share
Agriculture	599	25%
Other	352	15%
Government	297	13%
Research	284	12%
Environmental design / consulting	241	10%
Landscape management	217	9%
Water district	163	7%
Student	105	5%
Commercial / Industrial	51	2%
Golf course management	49	2%

Table 2.1: Respondent counts for the online survey.

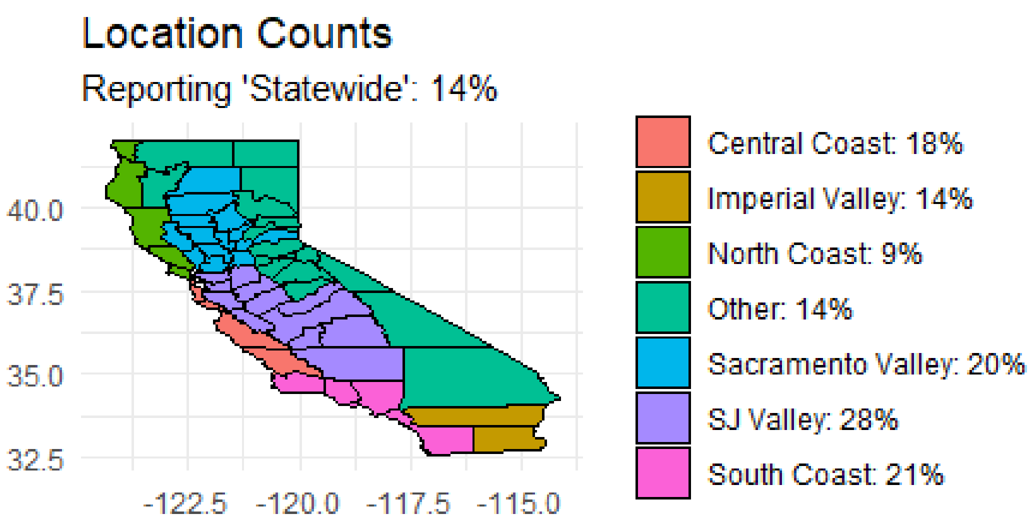

Figure 2.1: Counts of responses for location of CIMIS related activities.

shares decrease when moving in the "never" direction. On the other hand, one can observe an opposite trend for insolation (sun radiation), soil temperature, and relative humidity, which seem to be of less interest for respondents. Interestingly, air temperature seems less correlated with the frequency response, with response rate for "often" lower than "sometimes". This could stem from the use of air temperature data: while irrigation requires using ET data often, air temperature data applications might require less frequent data pulls.

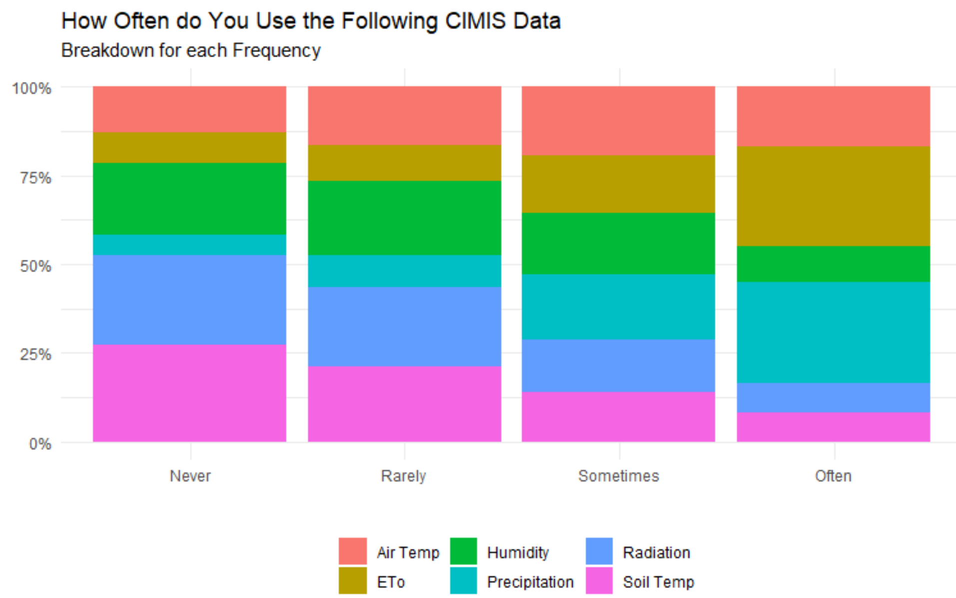

Figure 2.2: Response rates for frequency of each weather measure offered by CIMIS.

Respondents seem satisfied with CIMIS services. About 72% of respondents reported using CIMIS at least occasionally. The user types reporting "often" using CIMIS the most were Agriculture, followed by Golf Course Management and Water Districts. These user types are indeed likely to use CIMIS on a day to day basis, at least for some part of the year. In research and planning, on the other hand, one might use CIMIS to draw data only at an initial stage of a given task. In general terms, of the respondents who report using CIMIS to some extent, 77% say it is at least "moderately important" for their operations, with 22% reporting CIMIS as "extremely important". The frequency of use and importance scores are positively correlated: frequent users also report high importance of CIMIS to their operations, which makes sense. The correlations between frequency and satisfaction, and between importance and satisfaction, seem less pronounced. There might be users who use CIMIS infrequently, perhaps because only a smaller part of their tasks involve the weather or climate information provided. Nevertheless, they seem satisfied with CIMIS services, as the satisfaction scores are relatively high.

We also asked respondents to rank factors which hinder further use of CIMIS. Various answers were provided, given the results of initial surveys, and there was also room to specify other answers. Two main concerns exist, especially for users in agriculture: how reliable is the data and how to integrate it into existing systems and practices. Many growers and consultants in agriculture complement CIMIS with other data sources, such as soil moisture sensors, irrigation logs, and flow meters. Integrating information from multiple sources into decision making is a challenge faced by virtually all growers.

2.5 Economic gain assessments

Agriculture

599 respondents, about a quarter of our survey, reported agriculture to be their primary business. Out of these, about half (331) work on one farm, and the rest (267) are consultants of sorts (i.e. work on many farms). 89% of respondents in agriculture report using CIMIS to some extent. Growers and consultants were asked to report their total acreage, selecting into pre-determined ranges. Summing these, we have 318,156 acres covered by growers, and almost 3 million acres covered by consultants.

Many of the questions for growers and consultants were similar. One notable exception is regarding water use. The team decided not to ask growers how much water they use, fearing that growers would not want to share this information and would not finish the survey. However, consultants were asked how much water their clients use on average. This question was presented in the online survey as a slider bar, with a default at the lower bar (0.5 AF/acre), and an option to check a "Not applicable" box. This box was not checked very often. Instead, it seems like many consultants who did not want to answer this questions left the slider bar at the default value of 0.5 AF/acre. This is a very low value for irrigated crops, and we assume that all these responses are basically non-answers. Ignoring them, the average reported water use is 2.96 AF/acre per year (0.94 standard deviation, over 152 responses). This seems like a very reasonable distribution for water use in irrigated crops. Indeed, the USDA's most recent Farm and Ranch Irrigation Survey (2013, Table 4: "Estimated Quantity of Water Applied By Source: 2013 and 2008") reports a total of 7,543,928 irrigated acres in California, with a total of 23,488,939 AF of water applied, and a resulting average water use of 3.11 AF/acre, only a minor deviation of the reported average.

Given the responses from agricultural consultants, we seem to have captured a very large portion of the drip irrigated acres in California. As a baseline for valuation, we will use the total 2013 drip irrigated acreage from the USDA survey, 2.8 million acres. While some growers might use CIMIS with gravitational or sprinkler systems as well, our understanding of the qualitative and quantitative responses is that CIMIS is mostly important for drip. We exclude the potential of CIMIS values on non-drip acreage, noting that our estimates would therefore be conservative in that sense.

Water savings

Growers in our survey reported an average CIMIS water saving effect of 24.2%. The reported saving rates seem to be distributed evenly among crops and grower acreage. The average water saving rates reported for consultants is 21.5%, a slightly lower rate than the growers, but this difference is not meaningful in an economic or statistical way. Figure 2.3 plots the distributions of reported savings by growers and consultants, with very similar means and medians. Regressing the reported savings rate on all user types, one cannot reject the null hypothesis that the mean water saving effect is equal between growers and consultants with 95% confidence (see Table 2.2 for the results). Since each group deals with different acreages, we interpret this result as lack of substantial economies of scale in water saving by CIMIS.

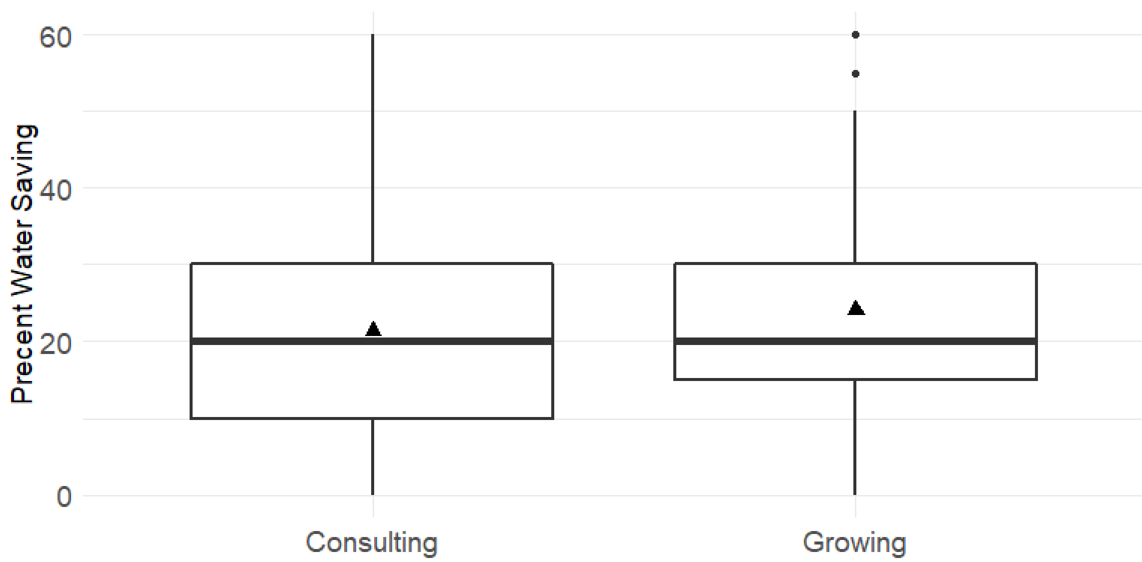

Figure 2.3: Box plots describing distribution of reported water savings with CIMIS. The box contains the 25-75 percentile observations. The bold horizontal line is the median. The triangles show the average for each group.

For a conservative estimate, we take the lower estimate as our representative saving effect. The average annual water use in the survey was 2.94 AF/acre, slightly lower than the California average in the USDA survey for all irrigated land. That same survey also reports that the average water use by drip irrigation is 2.5 AF/acre . We use this datum, which leads to more conservative results. Considering 2.8 million acres of drip irrigated land, with 2.5 AF/acre, the total water use in drip irrigation is 7 million AF per year. The (conservative) water saving rate of CIMIS is 21.5%, and this water use already incorporates it. Therefore, the amount of water which is saved by growers who use CIMIS is estimated at 1.92 million AF per year, or about 8.2% of the total water used for irrigation in California in 2013.

	Reported Water Savings (percent)
Consulting	-2.62^{*}
	$(-5.43, 0.19)$
Golf Course Management	-2.65
	$(-7.94, 2.64)$
Government	-2.70
	$(-6.42, 1.02)$
Landscape Management	6.07^{***}
	$(3.16, 8.97)$
Research	1.16
	$(-2.65, 4.97)$
Water District	-0.67
	$(-5.04, 3.70)$
Constant	24.08^{***}
	$(22.23, 25.93)$
Observations	692
Adjusted R^2	0.04
F Statistic	6.33^{***} (df = 6; 685)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 2.2: Regression of reported water saving percent with CIMIS on user categories. The excluded category (constant) are growers. The numbers in parentheses are the 95% confidence interval of the coefficients.

The "net" economic value of this water saved with CIMIS is the sum of expenditures that would have been incurred by growers if they were to purchase them. This assumes that this water could have been purchased (i.e. there would be enough water), that the price would be set at some level, and that the demand for water by growers is perfectly inelastic. The first and last assumptions seem reasonable, at least for an approximation. Assuming a water price, we can therefore multiply the amount of water saved by that price to get the "net" or intensive gains. However, water markets are not established enough in California to determine a single benchmark price for this calculation. In an evaluation of drip irrigation, Taylor, Parker, and Zilberman (2014) use a range of water prices to assess monetary gains from water savings. They use a lower price of $80/AF, high price of $220/AF, and an average price of $150/AF. However, during the drought in 2014, prices as high as $1,100 per AF were reported in the media (Vekshin, 2014). For the estimated 1.92 million AF of water saved with CIMIS, these prices imply monetary savings of $154 - $422 million, depending on the water availability which would determine the price. On a severe drought year, the saving could reach up to $2,112 million.

GDP gains from extra water in agriculture

The monetary cost of water saved can be viewed as savings on the intensive margin. One can also consider gains on an extensive margin. The water saved by use of CIMIS is likely to be used in agriculture as well. This means more acres can be grown with the same initial amount of water. The "full" economic value of the water saved by CIMIS in agriculture is the value of agricultural output that can be produced with it on acres not irrigated before.

This following analysis includes the economic value of growing alone, without the added values of post-harvest and economic multiplier effects, and probably a safe lower bound. We do not, however, include a counter-factual productivity of non-irrigated land. In California, this is probably range land or acreage that is too sloped for traditional irrigation methods, and therefore of very low economic productivity.

With 1.92 million AF of water saved by CIMIS, and an average use of 2.5 AF/acre by growers (assuming the water goes to drip irrigated crops), the savings from CIMIS can water an extra 768,000 acres in California. To put this in context, this is about double the total walnut acreage in 2016. Because of economic and technical constraints of water transport, it is hard to determine which crops would be planted in these extra acres. A conservative approximation assumes that the water saved by CIMIS serves to replicate the existing distribution of crops (rather than adding acres to the highest value crops), taking the average value of productivity of an acre as the benchmark. The weighted average of grower revenue per acre in 2016 (not including pasture and crops with less than 5,000 acres in total) was $3,757 per acre[1]. Multiplying by 768,000 acres, a conservative approximation for the contribution from CIMIS to California's GDP via agriculture is about $2.89 billion. This number may appear very high, yet this calculation took various conservative assumptions:

[1]Author's calculations based on crop report data by the California Department of Food and Agriculture.

in the calculation of the water saved, in assuming the value of extra acreage, and in not including post-harvest added value and multiplier effects. To be even more conservative, let us assume that the elasticity of demand for the products grown on these extra acres is -2. That is, an increase of 1% in quantity would drop the price by about 0.5%. This is a reasonable estimate for elasticities of high value crops (e.g. the elasticity of demand for pistachios is about -1.5). The resulting extra income for growers is then about $1.44 billion dollars.

Yield effects

CIMIS allows for more precise irrigation, which means not only saving water but also increasing yields: water application can be adjusted to the plant requirements, which might depend on the weather and growing phase. We ask growers and consultants how does CIMIS contribute in increasing yields, ranking from 1 ("None") to 5 ("A lot"). How should we quantify these ranked contributions? Taylor, Parker, and Zilberman (2014) mention average yield effects of drip irrigation, ranging between 5% and 25% increase in output. This extra yield effect is explained by allowing for more consistent soil humidity and the precision of the irrigation. This aspect of drip depends on weather and ET information, such as the one provided by CIMIS, to assess the water intake by plants and the appropriate amount of water required. We calculate an average yield effect of CIMIS by reconciling the respondent rankings with a portion of the yield effects from drip irrigation. For a lower estimate, rankings between 1 and 3 ("Somewhat") are attributed 0% yield effect, and the rankings of 4 and 5 get 5%. For a higher estimate, ranking of 1 gets 0% yield increase, ranking of 2 and 3 get 5% yield increase, and the rankings of 4 and 5 get a 10% yield increase. These percent yield effects are then averaged among the respondents. We aggregate growers and consultants with equal weights. 41% of respondents rank the importance of CIMIS for yield effects at 4-5. The low estimate for yield contribution of CIMIS results in 2% output increase, and the higher estimate at 5.9% increase. At a conservative estimate of per-acre income of $3,757 for growers, this represents an extra yearly income of $76 - $222 per acre. For the 2.8 million acres using drip irrigation, this would account for $213 - $622 million yearly from the contribution of CIMIS to yields. Assuming again the demand is elastic with a coefficient of -2, these estimates would halve to $107 - $311 million.

Quality effects

Weather data can have quality effects on crops. For example, using ET data and drip irrigation, the quality of tomatoes (measured by sugar content) can be increased by controlled irrigation deficit in proper timing. For tomatoes grown under a contract, reaching threshold quality levels raises the price received by the grower (Ayars, Fulton, and Taylor, 2015). Another potential use of weather data is in pest control, avoiding not only yield loss but quality degradation as well. These two examples reflect a relationship between quality and price that has long been established in the literature (Parker and Zilberman, 1993). To assess the contribution of CIMIS to quality, we also asked respondents to rank it from 1 ("None") to 5 ("A lot"). We assume that a score of 4-5 represents a quality index resulting in a price increase of 5%. About 45% of all respondents (growers and consultants) report a score of 4-5. The average price increase due to quality is therefore 2.2%, or $83 per acre. For 2.8 million acres, this results in a total increased revenue of $231 million. Note that this price (and revenue) increase is due to quality improvement, and thus not accompanied by a quantity reduction in our analysis.

Landscaping and Golf

These are gains from water saving in parks, golf courses, and gardens. They were assessed as a small portion of the total gains from CIMIS in the 1996 report by Parker et al., totaling about $2.3 million (equivalent to about $3.8 million in 2019). Our current estimate for these gains is much higher. The discrepancy from the 1996 report is due to several factors. First, we believe to have reached out to more respondents in this sector. Second, water prices in California have gone up substantially. Third, there might be more use of CIMIS and smart irrigation planning in the sector compared to 20 years ago.

We focus on responses from landscape managers and golf course managers. They report their operating acreage (selecting one of several provided ranges), the average water use, and the estimated saving rate by using CIMIS. We have 28 respondents in golf courses with 6,750 acres in total, and 137 respondents in landscape management with 179,000 acres. The total sum is about 21 times the acreage of the equivalent category in the 1996 report.

Based on the initial interviews, we grouped them into a single user category, but still asked them to select into landscape or golf later in the survey. Table 2.2 proved us wrong. Surprisingly, it turned out that the users in landscape management reported much higher water saving rates with CIMIS. This could potentially be explained by technology: big turf areas are still likely to be irrigated with sprinklers, which allow lower savings rates even if CIMIS is used for optimal water calculations. On the other hand, a lot of non-turf landscaping might be irrigated with drip.

The total amount of water, saved yearly with CIMIS according to our respondents, is 220,707 AF. Water prices for these types of users are much higher than in agriculture. We can use the municipal water rates to get an estimate of the monetary savings. The EBMUD rates, effective 2018, are $5.29 per 100 cubic feet ("All Other Accounts" – nonresidential) or $4.12 for non-potable water. The Los Angeles Department of Water and Power charges commercial, industrial and governmental users by tiers. For January 2019, the tier 1 rates are $5.264 per 100CF, and tier 2 rates are $8.667. The specific tier 1 allotment is set for each user. However, some non-profit users might get rates as low as $2.095 for tier 1 and $3.595 for tier 2. For comparison with agriculture, note that the lowest rate cited above for municipal water is more than four times higher than the "high" rate for agriculture in Taylor, Parker, and Zilberman (2014). The spread of prices, even within municipalities, suggests that they might not reflect the marginal cost of providing water to consumers. However, water utilities (public and private) have regulated rates and usually work on a "cost plus" basis, such that the water rates should reflect their real average cost. These rates can therefore be used to assess the economic gains from water savings.

The different municipal rates serve to construct bounds for our estimates. This first order approximation does not take into account the potential elasticity in water demand, or the potential effect of CIMIS in lowering residential water pricing by curbing down demand. However, we think they are good benchmarks and could definitely serve as an estimate for order of magnitude. The lower rate is the LADWP non-profit rate, which might not apply for many CIMIS users. Assuming nobody exceeds their tier 1 allocation, the value of water savings amounts to $201 million per year. For a higher EBMUD rate of $5.29, the savings amount to $509 million per year. For a reasonable upper bound, assuming we are in Los Angeles and 90% of the water consumption is in tier 1 (i.e. we exceed the initial allotment by 10%), the sum is $539 million.

Unlike the case of agriculture, we do not believe the survey responses in this category have captured all (or nearly all) the relevant acreage. Neither do we have a good sense of the total relevant acreage in California, which could indicate by what factor these estimated gains could be extrapolated. However, the sums are substantial as they are. We take them as our total estimates for gains from CIMIS, noting that they are an under-estimate in this sense.

2.6 Discussion and conclusions

This chapter analyzes the gains from CIMIS, focusing on agriculture and some urban uses. The gains are much higher than the ones found in the 1996 report. This is partly due to increased economic activity in general, but probably has to do with more adoption of smart irrigation as well. The total yearly gains in agriculture range between $492 million, taking only the intensive margin effects, and up to about $1,982 million considering the extra acres that can be grown with the saved water. A surprisingly large sector using CIMIS is landscaping and golf courses, with yearly monetary savings of at least $201 million for our survey sample alone.

Several other user types were included in the survey, indicating a substantial role of CIMIS in areas crucial for California's economy. Respondents use CIMIS to plan drainage in agricultural and urban settings, taking advantage of CIMIS historic rainfall records. CIMIS is used for water budgeting and even pricing. Researchers in the public and private sector use CIMIS for diverse purposes, from basic research to calibration and verification of other weather related products. These are just a few of many additional uses of CIMIS we know about, but do not quantify here due to the complicated methodological framework required.

The economic gains from CIMIS surely surpass the ongoing costs of a system with less than a dozen employees. However, could these gains be achieved by the private sector? The decreasing costs of weather sensors mean that growers and other users could potentially access precise data on their own. If we wanted a cheap weather station, costing about $1,000, for every 1,000 acres of drip irrigated land in California, the total cost would surpass

$2.8 million, plus some ongoing costs for maintenance. This, however, would prevent many benefits from the centralized aggregation of data and the historical records that are crucial for research and planning, as one could not assure that aggregation of the data from all these separate private stations would occur. While several online aggregators of weather information exist, many rely on the public information provided by networks such as CIMIS and other government bodies such as airports and air quality monitoring systems. It is not obvious that private aggregators would be profitable if they had to purchase this information, or what their WTP would be. Moreover, the ET measurements which many growers use are usually not captured by commercial stations, and there are concerns regarding the reliability of ET approximations by other variables. The development of satellite technology might change these conditions in the future.

Chapter 3

Estimating The Impact Of Heat On Pistachio Yields: Small Panel Meets Big Data

3.1 The challenge with a perennial crop

This chapter deals with the non-linear effects of temperature on yields in California pistachios (*Pistachia vera*). Dealing with a perennial crop presents challenges rarely encountered in research on crops such as corn, soybeans, and rice. A first challenge has to do with biology. Annuals are grown from seed every year. Each yield observation is treated as an independent draw from a distribution. Some adjustments for spatial correlation might be taken, but the interactions of consecutive yields, e.g. via pest build-up or changes in soil chemistry, are mostly overlooked.[1] Perennials, on the other hand, do experience yield effects of factors such as tree vintage, carry-over from past years, and alternate bearing patterns. These processes are responsible for some of the factors obfuscating the real relationships between temperatures and nut yields (Pope et al., 2015). Altogether, the potential for statistical noise, stemming from correlations between error terms, is much higher in perennials than annuals.

A second challenge is that temperature might affect yields not only in the growing season and not only via the common implicit "heat stress" mechanism. For perennials, temperature effects might be greatest before or after fruit bearing time. This chapter deals with the effects of temperature on pistachios during their winter dormancy phase. The following brief explanation of dormancy is based on Erez (2000). Many fruit and nut trees, including pistachios, have a dormancy phase during winter. This phase is an evolutionary adaptation, allowing the tree to "hibernate" and protect sensitive organs while harsh weather conditions take place. Trees prepare for dormancy by storing energy reserves, shedding leaves, and developing organs to protect the tree buds. Once a tree went into dormancy, it needs to

[1] This might not be a bad approximation with the common application of pesticide and fertilizer.

calculate when to optimally "wake up". Blooming too early might expose the foliage to frost. Blooming too late means not taking advantage of available resources (sunlight), and eventually being out-competed. Trees use environmental signals to trigger bud breaking and bloom. These signals involve day length and temperatures. Failure to attain threshold signal levels, varying between crops and varieties, leads to late, low, and non-uniform bud breaking, which is linked to low yields at harvest. This threshold mechanism means that small changes in the temperature distribution can have large effects on yields, especially in the warmer areas where the chances of not reaching the threshold signal are higher.

Several agronomic models exist for this dormancy exit mechanism and the role of temperature in it. The Dynamic Model (see Fishman, Erez, and Couvillon, 1987b,a; Erez et al., 1989) seems to be the most precise in predicting bloom in many temperate areas such as California (Luedeling, 2012). This model uses a metric of chill portions (CP), which are calculated with a vector of hourly temperatures. The formula is sequential, mimicking chemical dynamics which depend on the concentrations of substrate and product. Chill portion build up depends on these concentrations and the ambient temperature. Roughly speaking, when temperatures go above 6^oC, accumulation slows down. When temperatures exceed 15^oC, the process reverses, and the CP count quickly drops to the last integer portion that has been "banked". Thus, rising winter daytime temperatures can have a detrimental effect on chill count, even if the temperatures themselves are not extreme on the yearly distribution, because they interfere with the build-up of chill portions. This mechanism is an example of the complex modeling issues, from the biology perspective, when dealing with perennials: it require crop specific agronomic knowledge, and the CP build up model makes it impossible to assess the marginal effects of certain temperatures, as CP are not linear in degree hours.

A third challenge with some perennial crops is the limited information on yields. Heterogeneity in local weather conditions increases the statistical power of traditional yield panels in annuals, with acreage spreading over many geographic regions. California pistachios, on the other hand, are concentrated in the southern part of San Joaquin Valley. Moreover, they are planted in areas where the climatic conditions are mostly beneficial for them. Few events of adverse weather exist on record, which can be used for analysis. Therefore, the variance in CP in our range of interest is even more limited.

The issue of limited information also has to do with the size of reporting units in the available data. The California Department of Food and Agriculture, as well as the US Department of Agriculture, usually report average yields on the county level. If the counties are large, compared to the growing area, few observations will be generated, and the averaging process will get rid of useful extreme observations on the sub-county level. The aggregated reporting problem, together with crop concentration, limits the possibilities of traditional econometric analysis on crop yields. I address this problem here for California pistachios, but the challenge might prove a barrier for research on other crops as well. Consider not only high value commercial crops concentrated in a few California counties, but also "orphan crops": local crops which have received less attention from researchers and the private sector, yet generate substantial nutritional value for low income communities in developing countries. The African Orphan Crops Consortium, an initiative to promote research and use of these crops in Africa, list 101 crop of interest on its website, many of them perennial.[2] Cullis and Kunert (2017) note that orphan crops "...are poorly documented as to their cultivation and use, and are adapted to specific agro-ecological niches and marginal land with weak or no formal seed supply systems". Research on specific orphan varieties might therefore suffer from the same challenges of California pistachios: biological complexity, concentration of growing acreage, and few data reporting units.

In this chapter, I combine two approaches to estimate the yield response of California pistachios to winter CP count. The first approach is a "big data" one: I enhance a California yield panel of five counties with local temperatures at the pistachio growing areas. I use satellite data and temperature readings from local weather stations to create a large data set that can be connected with the yearly yields. Substantially increasing the number of explanatory variables, this allows for more nuances observations. The second approach is an aggregate estimation methodology, previously used in agricultural productivity literature but –to my knowledge– not yet explored in climate literature. This approach notes that the (aggregated) observed outcome variable is a mix of unobserved sub-unit heterogeneity in the data generating process. Information about this heterogeneity is used to recover the relationship between temperatures and yields.

The result of this exercise is the first successful recovery of the nonlinear yield response to winter chill in commercial pistachio production. I apply my findings to climate predictions in the current growing areas to show the potential impact of climate change on California pistachios in the next 20 years, and predict that a significant decline can be expected.

3.2 Recovering Information With a Small Panel

The challenge is recovering information on the yield response to winter CP with a small panel. Just for comparison, Barrows, Sexton, and Zilberman (2014) estimate the yield effect of genetically engineered (GE) traits (two parameters) for cotton in a panel of 1,900 observations. Schlenker and Roberts (2009) have a main specification for corn with 16 temperature parameters (intervals of 3^oC) on a US county panel with 105,981 observations. Both studies also include a few more parameters and fixed effects which take more degrees of freedom, but still allow for meaningful estimations. My California pistachio panel has 51 CP measures and 165 observations (5 counties over 33 years).

Suppose we actually had sub-county yield data for pistachios (these could be individual orchards), so we could match a yield observation to each local temperature observation. Following the pest control literature (e.g. Zilberman et al., 1991), the yield can be modeled as a potential yield (PY) times a function of the temperature and an error term.

$$Yield_{i,t} = f(T_{i,t}) \cdot PY_{i,t} \cdot e^{\varepsilon_{i,t}} \tag{3.1}$$

$$log\,(Yield_{i,t}) = log\,(f(T_{i,t})) + log\,(PY_{i,t}) + \varepsilon_{i,t} \tag{3.2}$$

[2] http://africanorphancrops.org/meet-the-crops/

The potential yield is the (average) yield that would have been attained with zero damage temperatures. Note, this is not the maximum yield, but the average yield that would be attained with zero damage stemming from our input of interest. This does allow this average yield to be lower than an "ideal" yield, as the crop might experience sub-optimal levels of other inputs even when the damage from pests or temperature is zero. In fact, the setup assumes that damages from temperature are orthogonal to damages from other factors (e.g. alternate bearing, pests, etc.).

The log form, which also approximates the effect of temperature as percent change in yield, allows us to separate the temperature effect from background potential yield and noise, and estimate it by OLS or similar methods. What happens if we only have the aggregated yields on the county level? Barrows, Sexton, and Zilberman (2014) deal with a similar question when estimating the yield effect GE varieties in cotton, corn, and soybeans. They have a panel of country level yields, but these are not partitioned to GE and non-GE yields. However, they do have the shares of GE and non-GE planted acres in each country. treating the total yields as weighted averages of the GE and non-GE yield, they can estimate the yield effects of GE traits using OLS.

The problem here is similar. I only have county level yields for pistachio, but I do have the shares of each county experiencing various CP levels. The total effect of temperature on the county level is the average of its effects on the individual orchards, weighed by their acreage share. This turns out model into:

$$log\left(Yield_{c,t}\right) = log\left(\sum_{CP=36}^{86} \Phi_{c,t}(CP) \cdot f(CP)\right) + log\left(PY_{c,t}\right) + \varepsilon_{c,t} \tag{3.3}$$

where CP is the winter CP count, ranging from 36 to 86 in my data, and $\Phi_{c,t}(CP)$ is the *share* of county c in year t that experienced that CP count. These are my weights. Ideally, I could estimate the first RHS expression by a non-parametric binned regression, and get the average treatment effect of each CP level. However, in a small panel setting and with numerous CP values, there might not be enough statistical power to estimate an equation such as:

$$log\left(Yield_{i,t}\right) = \sum_{CP=36}^{86} \Phi(CP)_{i,t}\boldsymbol{\beta})_{CP} + log\left(PY_{i,t}\right) + \varepsilon_{i,t} \tag{3.4}$$

with different parameters $\boldsymbol{\beta})_{CP}$ for each chill portion level. The high resolution in the CP share variables is necessary, however, if we want to determine the shape and location of a non-linear effect. I need to reduce the dimensionality of the problem without getting rid of my specific county shares at each CP. Two solutions are applied here: structural and polynomial.

Structural Estimation

The structural approach follows the original potential yield setting. When PY is viewed as the zero (relevant) loss yield, the function $f(\cdot)$ becomes a net-of-loss function, ranging between 0 (no yield remaining) and 1 (potential yield attained). In practice, however, the potential yield is unknown, and researchers will often use a fixed-effects setup to model the heterogeneity in potential yields between countries or regions. For example, Barrows, Sexton, and Zilberman (2014) motivate their model with an "expected efficacy" function that is bounded in $[0, 1]$ (page 680) but later set up and run a linear fixed-effects model (page 687). This econometric practice is very common in many settings, but since the potential yield is no longer modeled as an expected maximum, the function $f(\cdot)$ can no longer be interpreted as a net-of-loss damage function which is bounded in $[0, 1]$.

I suggest the following setup, which follows the original motivation more closely. The agronomic literature, introduced more thoroughly in a subsequent section of this chapter, looks at the yield response in pistachios as a satiated process. When the weather conditions are too warm, there would be virtually no yield. When the weather conditions are cold enough, yield would be normal (conditional on non-temperature factors). However, colder conditions will not have further yield effects (within the scope of California climate). I therefore take the potential yield to be the (county-decade) average yield, considering only years when chill is deemed as sufficient by the existing literature in the entire county. This would be equivalent to Barrows, Sexton, and Zilberman having a measure of pest infestation on non-GE areas, and using the country averages yields when the infestation level is relatively low as potential yield. For this, I take a CP level of 65, which has not been shown to reduce yields in previous publications (e.g. Pope et al., 2015). Of my 165 observations, 101 serve to calculate the potential yield, and 64 are not. The rate seems high, but it assures at least two observations are used per county-decade to calculate the potential yield.

Now, the ratio of yield to potential yield in the panel should theoretically be bounded between 0 (in very low chill) and 1 (when chill is optimal), with deviation from these bounds attributed fully to the disturbance term.[3] Assuming that this increase is smooth and monotonic, the logistic probability function is a good candidate to model the process. Equation (3.3) turns into:

$$log\left(\frac{Yield_{c,t}}{PY_{c,t}}\right) = log\left(\sum_{CP=36}^{86} \Phi_{c,t}(CP) \cdot P(CP \mid \boldsymbol{\delta})\right) + \varepsilon_{c,t} \tag{3.5}$$

where P is the logistic probability of CP given the parameters $\boldsymbol{\delta}$. The logistic distribution has two functional: location (which is also the mean and median of the distribution) and scale (a second moment parameter, in fact a multiple of the distribution variance). This reduces

[3]Theoretically, at low chill levels the error term is bounded below as the ratio cannot be negative. This is a potential violation of the zero conditional mean assumption (i.e. the expected error term might be positive for low chill values). It seems like less of a problem with my aggregated data, where the minimal ratio is 0.195. In theory, the bias would attenuate the chill effects, making my estimates conservative.

the dimensionality of the problem, as only these two parameters need to be estimated, resulting in more statistical power. These parameters are found by non-linear regression, minimizing the sum square errors:

$$\widehat{(l,s)} = arg\min_{l,s} \sum_{c,t} \left[log\left(\frac{Yield_{c,t}}{PY_{c,t}}\right) - log\left(\sum_{CP=36}^{86} \Phi_{c,t}(CP) \cdot \frac{1}{1+exp(\frac{l-CP}{s})}\right)\right]^2 \quad (3.6)$$

The numerical solution, minimizing the sum squared residuals in the model, is run in R using the “nls.lm” function of the “minpack.lm” package (Elzhov et al., 2016).

Polynomial Regression

This approach follow the more standard procedure with fixed effects, where the yield effect of temperature might not be bounded above by 1. The interpretation of the estimated effect is the change in yield from the average growing condition. To deal with the dimensionality challenge of equation (3.4), I approximate the effect of chill on log yields using a Chebyshev polynomial, one of the approaches used by Schlenker and Roberts (2009) in their corn study. The equation is:

$$log\,(Yield_{s,t}) = \mathbf{\Phi}'_{s,t} \sum_{j=0}^{2} \gamma_j \cdot \mathbf{U(j)} + log\,(PY_{s,t}) + \varepsilon_{s,t} \quad (3.7)$$

where $\mathbf{U(j)}$ are vectors of the polynomial components values at each CP in $[36, 86]$, and γ_j are the coefficients to estimate. I center both the response and shares at each CP around their averages to account for county-decade fixed effects. The coefficients are found numerically using the same R package.

3.3 Pistachios and Winter Temperatures

California pistachios are a high value crop, with grower revenues of $1.8 billion in 2016. The most common variety is “Kerman” (with “Peters” for male trees), and almost all the California acreage is planted in five adjacent counties in the southern part of the San Joaquin valley. In recent years, rising winter daytime temperatures and decreasing fog incidence have lowered winter CP counts. Climatologists have concluded that winter chill counts will continue to dwindle (Luedeling, Zhang, and Girvetz, 2009; Baldocchi and Wong, 2008), putting pistachios in danger at their current locations.

To better predict the trajectory for this crop and make informed investment and policy decisions, the yield response function to chill must first be assessed. This task has proven quite challenging. The effects of chill thresholds on bloom can be explored in controlled environments, but for various reasons these relationships are not necessarily reflected in commercial yield data. For example, Pope et al. (2015) report that the threshold level of CP for successful bud breaking in California pistachios was experimentally assessed at 69, but could not identify a negative response of commercial yields to chill portions of the same level or even lower. They use a similar yield panel of California counties, but only have one "representative" CP measure per county-year. Using Bayesian methodologies, they fail to find a threshold CP level for pistachios, and reach the conclusion that "Without more data points at the low amounts of chill, it is difficult to estimate the minimum-chill accumulation necessary for average yield."

The statistical problem of low variation in treatment (effective weather) at the growing area, encountered by Pope et al., is very common in published articles on pistachios. Simply put, pistachios are not planted in areas with adverse climate. Too few "bad" years are therefore available for researchers to work with when trying to estimate commercial yield responses. An ideal experiment would randomize a chill treatment over entire orchards, but that is not possible. Researchers resort either to small scale experimental settings, with limitations as mentioned above, or to yield panels, which usually are small in size (i.e. small number of yield reporting units), length (in years), or both.

Zhang and Taylor (2011) investigate the effect of chill portions on bloom and yields in two pistachio growing areas in Australia, growing the "Sirora" variety. Using data from "selected orchards" over five years, they note that on two years where where chill was below 59 portions in one of the locations, bloom was uneven. Yields were observed, and while no statistical inference was made on them, the authors noted that "factors other than biennial bearing influence yield". Elloumi et al. (2013) Investigate responses to chill in Tunisia, where the "Mateur" variety is grown. They find highly non-linear effects of chill on yields, but this stems from one observation with a very low chill count. Standard errors are not provided, and the threshold and behavior around it are not really identified. Kallsen (2017) uses a panel of California orchards, with various temperature measures and other control variables to find a model which best fits the data. Unfortunately, only 3 orchards are included in this study, and the statistical approach mixes a prediction exercise with the estimation goal, potentially sacrificing the latter for the former. Besides the potential over-fitting using this technique, the dependent variables in the model are not chill portions but temperature hour counts with very few degree levels considered, and no confidence interval is presented. Finally, Benmoussa et al. (2017) use data collected at an experimental orchard in Tunisia with several pistachio varieties. They reach an estimate for the critical chill for bloom, and find a positive correlation between chill and tree yields, with zero yield following winters with very low chill counts. However, they also have many observation with zero or near-zero yields above their estimated threshold, and the external validity of findings from an experimental plot to commercial orchards is not obvious.

Data

Pistachio growing areas are identified using USDA satellite data (Boryan, Yang, and Willis, 2014) with pixel size of roughly 30 meters. About 30% of pixels identified as pistachios are singular. As pistachios don't grow in the wild in California, these are probably miss-identified pixels. Aggregating to 1km pixels, I keep those pixels with at least 20 acres of pistachios in them. Looking at the yearly satellite data between 2008-2017, I keep those 1km pixels with at least six positive pistachio identifications. These 2,165 pixels are the grid on which I do temperature interpolations and calculations.

Observed temperatures for 1984-2017 come from the California Irrigation Management Information System (CIMIS, 2018), a network of weather stations located in many counties in California, operated by the California Department of Water Resources. A total of 27 stations are located within 50km of my pistachio pixels. Missing values at these stations are imputed as the temperature at the closest available station plus the average difference between the stations at the week-hour window.

Future chill is calculated at the same interpolation points, with data from a CCSM4 model CEDA (2016). These predictions use an RCP8.5 scenario. This scenario assumes a global mean surface temperature increase of $2^o C$ between 2046-2065 (from a baseline of 1986-2005) (IPCC, 2013). The data are available with predictions starting in 2006, and include daily maximum and minimum on a 0.94 degree latitude by 1.25 degree longitude grid. Hourly temperature are calculated from the predicted daily extremes, using the latitude and date (procedure coded in R by Luedeling, 2017). I then calibrate these future predictions with quantile calibration procedure (Leard and Roth, 2019), using a week-hour window.

Past observed and future predicted hourly temperatures in the dormancy season (November 1st to March 1st) are interpolated at each of the 2,165 pixels, and chill portions are calculated from these temperatures. Erez and Fishman (1997) produced an Excel spreadsheet for chill calculations, which I obtain from the University of California division of Agriculture and Natural Resources, together with instructions for growers (Glozer, 2016). For speed, I code them in an R function (Trilnick, 2018).

The data above are used for estimation and later for prediction of future chill effects. For the estimation part, I have a yield panel with 165 county-year observations. For each year in the panel, I calculate the share of county pixels that had each CP level. For example: in 2016, Fresno county had 0.4% of its pistachio pixels experiencing 61 CP, 1.8% experiencing 62 CP, 12% experiencing 63 CP, and so on. The support of CP through the panel is $[36, 86]$. Past county yields are from crop reports published by the California Department of Food and Agriculture.

Figure 3.1 presents chill counts and their estimated effects in percent yield change for two time periods: 2000-2018 and 2020-2040. The top left panel shows the chill counts in the 1/4 warmest years between 2000 and 2018 (observed temperatures). The top right panel shows the chill counts in the 1/4 warmest years in climate predictions between 2020 and 2040. Chill at the pistachio growing areas is likely to drop substantially within the lifespan of existing trees.

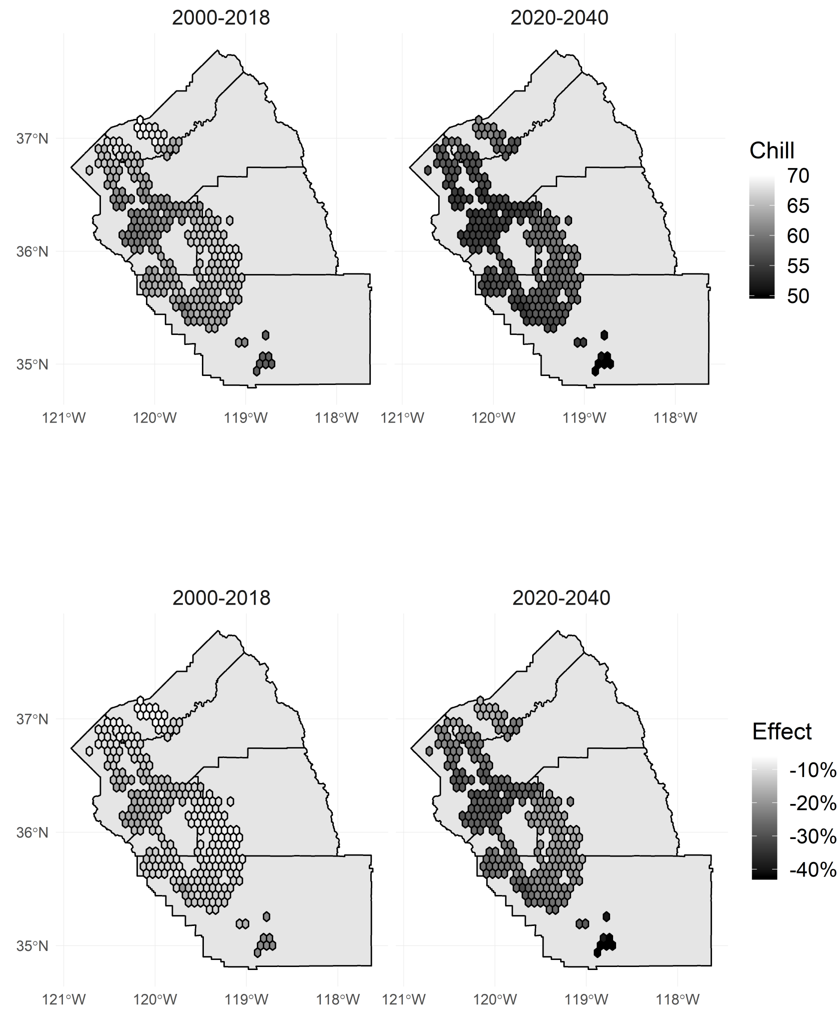

Figure 3.1: Chill and yield effects in the five main pistachio growing counties of California. This figure shows data and predictions for the first quartile chill portion measure, i.e. the warmest 1/4 years. Chill data for 2000-2018 is observed, and data for 2020-2040 is predicted. The yield effect is calculated with the logistic function fitted to the data.

3.4 Results

Results from the structural estimation are presented in Table 3.1. Both estimates for the parameters of the logistic function are statistically significant. The point estimates for location and scale are used to construct a logistic curve serving as my main functional estimation. In Figure 3.2, the bold smooth line shows the point estimate curve, and the shaded area is the 90% confidence area. Both are shifted down by one unit to portray the negative effects compared to a high chill benchmark. The results show a visible yield decline when CP go below 70, which is consistent with the experimental literature reporting bloom inconsistencies below 69 CP. At 55 CP, which have only been experienced locally so far, the loss is about 30%.

```
Parameters:
     Estimate Std. Error t value Pr(>|t|)
[1,]   47.430      4.085  11.612  < 2e-16 ***
[2,]    8.202      2.546   3.221  0.00154 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3805 on 163 degrees of freedom
Number of iterations to termination: 7
Reason for termination: Relative error in the sum of squares is at most ‘ftol’.
```

Table 3.1: Output from structural model.

Results from the polynomial regression are presented in Table 3.2 . The first coefficient is for an intercept term, and it is a zero with very wide error margins. This makes sense, as centering around the means also gets rid of intercepts. The second coefficient (for the Chebyshev component x) is positive, as we would expect, and statistically significant. The third coefficient (for the component $2x^2 + 1$) is negative, as we would also expect since the returns from chill should decrease at some point, but not statistically significant even at the 10% level. However, as dropping it would eliminate the decreasing returns feature, I keep it at the cost of having a wide confidence area. With the estimated coefficients, I build the polynomial curve that represents the effect of temperatures on yields. It is presented in Figure 3.2 with a bold dashed line. The 90% confidence area boundaries are the dotted lines bounding it above and below. Note that the upper bound of the confidence area does not curve down like the lower one. This is the manifestation of the third coefficient's P-value being greater than 0.1.

In both cases, the confidence area was calculated by bootstrapping. The data was resampled and estimated 500 times, producing 500 curves with the resulting parameters. At each CP level, I take the 5^{th} and 95^{th} percentiles of bootstrapped curve values as the bounds

```
Parameters:
       Estimate Std. Error t value Pr(>|t|)
[1,]  0.000e+00  1.062e+07   0.000  1.00000
[2,]  5.698e-01  1.841e-01   3.095  0.00232 **
[3,] -2.216e-01  1.474e-01  -1.503  0.13471
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3617 on 162 degrees of freedom
Number of iterations to termination: 2
Reason for termination: Relative error in the sum of squares is at most ‘ftol’.
```

Table 3.2: Output from polynomial model.

for the confidence area. This approach also deals with the potential spatial (or other) correlation in error terms. Another minor issue requiring the bootstrap approach is that the implicit potential yield estimation (by the “within” transformation or manually averaging yields of high chill years) should change the degrees of freedom in the non-linear regressions when estimating the standard errors.

In the lower panel of Figure 3.2, a histogram of positive shares is presented. That is, for each chill portion, the count of panel observations where the share of that chill portion was positive. The actual shares of the very low and very high portions are usually quite low. This shows the relatively small number of observations with low chill counts.

3.5 Discussion and conclusions

The two yield effects curves look very similar in the relevant chill range. By both estimates, the yield loss is very close to 0 at higher chill portions, and starts declining substantially somewhere in the upper 60’s, as the experimental literature would suggest. Interestingly, the polynomial curve does not exceed zero effect, although it is not mechanically bounded from above like the logistic curve. This probably reflects the fact that historically, the average growing conditions has not deviated much from the optimal range. The “within” transformation hence did not deviate the potential yield much from the optimum in this case.

At currently low chill portion ranges of 55-60, the effect is around -25%, again consistent with the stipulation of Pope et al. (2015) that a significant effect threshold would be located there. Considering alternate bearing and other factors contributing to the background fluctuation in yields, it is easy to understand how such effects on relatively small areas within the pistachio growing counties have not been picked up by researchers so far. Anecdotal yield losses due to low chill have happened on relatively small scale and passed undetected in the county-level statistics, especially when only one or two chill measures per county were

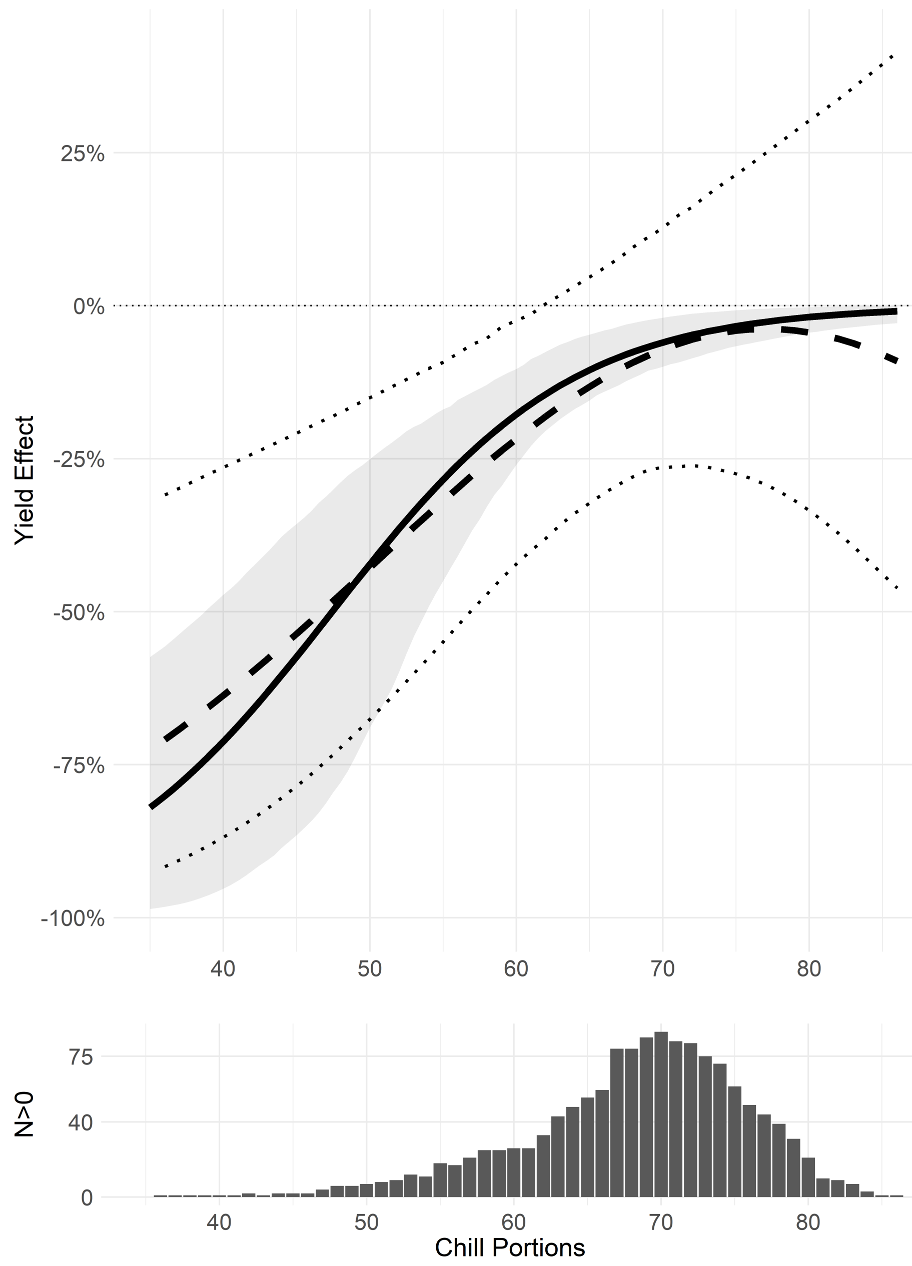

Figure 3.2: Effects of chill portions on pistachio yields. The solid bold curve and shaded area around it are the logistic fit with 90% confidence area (shifted down by unit to show negative effects). The bold dashed line and finer dotted lines are the polynomial fit and confidence area. The bottom panel shows the counts of observations with positive shares of each chill portion.

considered.

In this case, while the resulting curves are very similar, I find the structural approach more convincing. First, it has a smaller confidence area, and therefore seems more precise. Second, a polynomial of low order will not approximate the process described by agronomists very well. However, estimating higher order polynomials results in estimates that are not statistically significant.

The implications of my estimates for pistachio yields are depicted in the lower half of Figure 3.1. The bottom left panel shows the effects on the 1/4 warmest years in 2000–2018. They are mostly between 10-20% yield decline. These rates are easy to miss due to substantial yield fluctuations in pistachios. What do these estimates mean for the future of California pistachios? Prediction of yield effects for the years 2020–2040 are depicted in the bottom right panel, again for the 1/4 warmest years in the 2020-2040. They show substantial yield drops, which could amount to costs in the hundreds of millions of dollars. Chapter 4 in this dissertation explores the potential gains from a technology that could help deal with low chill in pistachios: applying kaolin clay mixtures on the dormant trees to block sunlight. Thee expected net present value of this technology is estimated at the billions of dollar in economic gains. Considering my results, there may be significant gains from using these technologies even in warmer years today.

Concluding this chapter, I want to stress the fact that even in the era of "big data" in agriculture, data availability is still a challenge when estimating yield responses to temperature in some crops, especially perennials and local varieties. Weather information required for assessing potential damages and new technologies might not always be available for a researcher. This chapter develops a methodology to recover this relationship, using local weather data and techniques for dealing with aggregated observations. I use this setup to empirically assess the yield effects of insufficient chill in pistachios, recovering this relationship from commercial yields for the first time in the literature. I then look at the threat of climate change to pistachio production in southern California. As winters get warmer, lowering chill portion levels are predicted to damage pistachio yields and disrupt a multi-billion dollar industry within the next 20 years. These results were made possible by using precise local weather data, applying relevant statistical methods, and using agronomic knowledge in the modeling process. This approach for information recovery from a small yield panel, with limited useful variability at first sight, could be useful for other crops as well.

Chapter 4

Micro-Climate Engineering for Climate Change Adaptation in Agriculture: The Case of California Pistachios

4.1 Micro-Climate Engineering

In the introduction chapter, I discuss the nature of temperature challenges posed by climate change. The rising average temperatures, according to the empirical literature, might not be the major source of potential loss. Rather, it's the elongating and fattening (right side) temperature distribution tails that would be responsible for much of the damage.

Could there be a way for farmers to target these tails directly? If so, such technologies could have potential uses for climate change adaptation. It so happens that farmers already deal with temperature extremes, and are capable of tweaking the tails of temperature distributions to avoid losses. The introduction already discussed "air disturbance technology", basically large wind generators, used to deal with some types of frosts (Hu et al., 2018).

Solutions for right side temperature tails exist as well. Of course, shading plants using nets or fabric is an existing practice, but these technologies are costly and not very flexible. However, other products that reflect sunlight and lower plant exposure to excess heat are available on the market. Perhaps the most common ones are based on a fine kaolin clay powder, which is mixed with water and sprayed directly on plants to form a reflective coat, sometimes referred to as a "particle film". These products have been commercially available since 1999, and are shown to effectively lower high temperature damages by literally keeping plants cooler (Sharma, Reddy, and Datta, 2015). Some manufacturers report a canopy temperature reduction of up to 6^oC when using their products. Spraying of this mix requires special rigs and equipment, but the costs are reasonable, and far lower than setting up shading in the form of nets (with or without frames). This technology can be thought of as cheap, disposable shading. Surprisingly, even though kaolin clay has been used by farmers to deal with other problems, less related to climate change (e.g. sunburns on fruit), I could find no economic literature discussing this technology. As with the case of air disturbance technology, these types of technologies have mostly been ignored by economists.

One reason for this gap in the literature could be that economists have not yet realized that these individual products and practices share a common conceptual framework: they are tweaking temperature distribution tails, while leaving the main probability mass untouched. This is an approach I call "Micro-Climate Engineering" (MCE). These are relatively small interventions in temperature distributions, limited in space and time, which aim to avoid the nonlinear effects of the extremes. Farmers know the available technologies for MCE and use them regularly, but their potential applications for climate change have not really been explored. The concept of MCE could be very important for climate change adaptation in agriculture, especially when considering the role of extreme temperatures on predicted future losses. MCE solutions, where feasible and profitable, could assist in preserving current crop yields and delaying more costly adaptation strategies.

This chapter sets to explore the concept of MCE in general, and assess the gains from MCE in California pistachios as a case study. Specifically, pistachios are threatened by warming winter days, which could threaten existing acreage within the next twenty years (see Chapter 3 for details). This challenge stands out in the existing literature in three ways: first, while much of the climate change literature focuses on annual crops, pistachios are perennial. This means that the opportunity cost of variety switching are higher. Second, the challenge does not occur in the "growing season", but on the winter months when trees are dormant and seemingly inactive. This emphasizes the importance of climate change effects year round, rather than just in the spring and summer. Third, the challenge stems from a biological mechanism that is not heat stress. Heat stress is perhaps the most obvious process by which rising temperatures can have adverse effects on yields, and by far the most studied in the economic literature on climate change. However, other biological mechanism are affected by weather as well, and can cause substantial yield losses. This paper incorporates agronomic knowledge on bloom disruption due to increased winter temperatures, a mechanism that is relatively unexplored in the economic literature.

Scientists at the University of California Cooperative Extension have been experimenting with kaolin clay applications on pistachios, and the results seem promising (Doll, 2015; Beede and Doll, 2016). This could mean a great deal to growers and consumers. This chapter analyzes the potential economic gains from this MCE application in California pistachios.

4.2 California Pistachios And Climate Change

Introduced to California more than 80 years ago, and grown commercially since the mid 1970's, pistachio (*Pistachia Vera*) was the state's 8th leading agricultural product in gross value in 2016, generating a total revenue of $1.82 billion dollars. According to the California Department of Food and Agriculture (2017), California produces virtually all pistachio in the

US, and competes internationally with Iran and Turkey (2/3 of revenues are from export). In 2016, five California counties were responsible for a 97% of the state's pistachio crop: Kern (35%), Fresno (28%), Tulare (15%), Madera (11%), and Kings (8%). Since the year 2000, the total harvested acres in these counties have been increasing by roughly 10% yearly. Each increase represent a 6 - 7 year old investment decision, as trees need to mature before commercial harvest (CPRB, 2009).

The challenge for California pistachios has to do with their winter dormancy and the temperature signals required for spring bloom. I discuss the dormancy challenge and the Chill Portion (CP) metric in Chapter 3. It is worth noting that in fact, for the areas covered in this study, chill portions are strongly (and negatively) correlated with the 90^{th} temperature percentile (Q90) between November and February, the dormancy season for pistachios. The correlation is very strong, with a goodness of fit rating of about 0.91. In essence, insufficient chill is a right side temperature tail effect, comparable with similar effects in the climate change literature.

Chapter 3 estimates the yield response of pistachios to CP. Substantial losses are predicted below 60 CP. Compared to other popular fruit and nut crops in the state, this is a high threshold (Pope, 2015), putting pistachio on the verge of not attaining its chill requirements in some California counties. In fact, there is evidence of low chill already hurting yields (Pope et al., 2015; Doll, 2015). Declining chill is therefore considered a threat to California pistachios.

Climate and Damage Predictions

Chill in most of California has been declining in the past decades, and is predicted to decline further in the future. Luedeling, Zhang, and Girvetz (2009) estimate the potential chill drop for the southern part of San Joaquin valley, where virtually all of California pistachio is currently grown. For the measure of first decile, i.e. the amount of CP attained in 90% of years, they predict a drop from an estimate of 64.3 ($\pm$2.9) chill portions in the year 2000 to estimates ranging between 50.6 ($\pm$3.8) and 54.5 ($\pm$3.6) (climate change scenario depending) in the years 2045-2060. Agronomists and stakeholders in California pistachios recognize this as a threat to this valuable crop (Doll, 2017; Jarvis-Shean, 2017). Together with increasing air temperatures, a drastic drop in winter fog incidence in the Central Valley has also been observed. This increases tree bud exposure to direct solar radiation, raising their temperature even further (Baldocchi and Waller, 2014). The estimates cited above virtually cover the entire pistachio growing region, and the first decile metric is less useful for a thorough analysis of pistachios. I therefore need to create and use a more detailed dataset, in fact the same one described in Cahpter 3. Figure 3.1 shows the geographic distribution of chill and potential damage in the 1/4 warmest years of observed climate (winters of 2000-2018) and predicted climate (2020-2040). While not very substantial in the past, these losses are predicted to reach up to 50% in some regions in the future.

4.3 Modeling Micro-Climate Engineering

This section develops a model to assess the gains from MCE. This is a single year, short run market model, solving for price and quantity under different winter chill realizations. Equilibrium price and quantity are used to calculate welfare outcomes such as grower profits, consumer surplus, and the total welfare. For each realization, the model is solved twice: once with an option to use MCE, and one without it. The differences in welfare outcomes under the same conditions, with and without MCE, are the welfare gains from MCE. Note that in both cases, agents act optimally. MCE gains are therefore to be interpreted as the difference in welfare measures between a world with MCE and a world without it.

I abstract from a benchmark with increased storage, which could theoretically alleviate inter-year fluctuations. Pistachios are usually stored for up to one year (Thompson and Adel A, 2016). The potential loss rates in a bad weather year are significant. Coping by storage in a meaningful way would require multi-year, double digit storage rate, which seems technically unfeasible.

Growers

The individual grower model draws from the pest control literature (see for example Lichtenberg and Zilberman 1986; Chambers and Lichtenberg 1994; Sexton et al. 2007; Waterfield and Zilberman 2012). In fact, the same setup is used to estimate the yield loss-rate from warm winters in Chapter 3 (the structural approach). Growers are considered to be small, facing the same prices for inputs and outputs, risk neutral, and fully informed about the prices the weather on their plot. Consider a grower with a production function $H(z)$, increasing in input z. The function $H(z)$ is referred to as the *potential output* function, where z is an input vector unrelated to the potential weather damage.

The grower also faces a damage or loss function $L(c) \in [0, 1]$. This loss depends on the chill realization this grower sees. The grower knows c before making input decisions z. This is especially realistic in our case, considering that most inputs (water, fertilizer, pest management, labor) are applied in the spring and summer, after the trees exit dormancy. The grower maximizes profits, manipulating the input level z.

Supply without MCE

A grower without MCE takes the weather related climate loss as exogenous, and maximizes profits by choosing an optimal level of input z:

$$\max_{\mathbf{z}} \ \pi = p \cdot [1 - L(c)] \cdot H(\mathbf{z}) - \mathbf{p_z}^T \cdot \mathbf{z} \tag{4.1}$$

Without loss of generality, treat z as a single, aggregate input. Note that the weather related loss is exogenous and constant. The grower's problem is solved by equating the value of marginal productivity of z to its price:

$$p \cdot [1 - L(c)] \cdot H_z(z) = p_z \tag{4.2}$$

The first order condition is solved for an optimal z^*, and the grower supplied quantity can then be calculated. The potential output function is specified as: $H(z) = \alpha + \beta \cdot \sqrt{z}$, which results in linear potential supply for the grower. Linear supply has traditionally been used in multiple settings, including on the returns from agricultural R&D (Norton and Davis, 1981; Alston et al., 2009)[1]. Together with the damage component, the supply function is as follows:

$$q(p,\, c) = [1 - L(c)] \cdot \left(\alpha + [1 - L(c)] \cdot \frac{\beta^2}{2p_z} \cdot p \right) \tag{4.3}$$

As expected, with higher price and/or higher net-of-loss rates, supply increases. To get realistic values, calibration of this function is required. Mainly, the coefficients of the potential output function need to be established. Available market data for calibration are historical county yields. I also have past weather data of 2,165 weather interpolation points of the same pixel size (1 Km^2), which I treat as individual growers. The linear supply form, however, can be aggregated as long as all bundled growers face the same weather realization. To model supply, I aggregate by the county chill quintiles. For each county-year, the 20% of interpolation points within the county with the lowest chill are the the first quintile; the next 20% chill points are the second, and so on. This assumes that, besides the temperature realizations, the potential output of all pixels (growers) in a county is identical. Note that for each county, the number of growers in each quintile is different, reflecting different acreage and capacity parameters. Altogether, the sum of county-chill quintile supplies should approximate the total supply.

To calibrate coefficients for county quintiles, I use market outcomes from 2016: the grower price and county quantities are taken from the California Department of Food and Agriculture annual crop report. To pinpoint a linear supply function, I also need a slope for supply, and use a short run supply elasticity parameter to calculate it. Short run elasticity in agricultural goods is usually considered very low on the short run (Alston, Norton, and Pardey, 1995, p. 321), and the 6-7 year setup requirement for pistachios should place its elasticity on the lower end even within this category. Others have modeled pistachio supply as completely inelastic (e.g. Gray et al., 2005), yet I think it is more realistic to take a positive parameter, as inputs such as harvesting effort can surely change supply. Estimates for supply elasticity are hard to come by in the literature. For an approximation, Russo, Green, and Howitt (2008) estimate the elasticity of almond supply w.r.t. one year lagged own price to be 0.19.[2] I take this as a starting point for the pistachio own price short run supply elasticity and use it in the main specifications. I later show results with other elasticities as well.

[1] A more recent review on the impact of biofuel demand on commodity prices matched estimates from the literature with estimates generated by a model with linear supply and demand at given elasticities. It turns out this simple model can give good approximations for the impact estimates generated by the more intricate models in the literature (see Persson, 2015).

[2] This estimate is not even statistically significant (p-value = 0.2), but it's the best I could find

With county quantities, market price, county-quintile chill related losses in 2016 (very low in all cases), and an elasticity, I can back out county coefficients for the supply function[3]. The county-quintile (c,k) coefficients are one fifth of the county coefficients, e.g. $\alpha_{c,k} = 0.2\alpha_c$. The total supply without MCE is the total sum of these county supplies:

$$Q(p) = \sum_c \sum_{k=1}^{5} q_{c,k}(p) = \sum_c \sum_{k=1}^{5} [1 - L(t_{c,k})] \cdot \left[\alpha_{c,k} + \frac{\beta_{c,k}^2}{2 \cdot p_z} \cdot (1 - L(t_{c,k})) \cdot p\right] \tag{4.4}$$

Supply with MCE

When MCE is available, the grower can also adjust the loss incurred due to weather. The profit maximizing problem is now:

$$\max_{x,z} \; \pi = p \cdot [1 - L(x,\, c)] \cdot H(z) - p_z \cdot z - p_x \cdot x \tag{4.5}$$

where x is the MCE input. Note that the natural chill itself, c, is still exogenous. This formulation assumes separability in output between x and z, i.e. input x only affects yields through the chill mechanism. Although some MCE products also have some other useful properties (e.g. some pest control capabilities and lowering water requirements), these properties are not very useful at time of tree dormancy. Hence, this assumption seems reasonable in our case.

An internal solution for the grower problem can be found with the two first order conditions, equating the value of marginal productivity of each input to its price:

$$p \cdot [1 - L(x)] \cdot H_z(z) = p_z \tag{4.6}$$
$$p \cdot [L_x(x)] \cdot H(z) = p_x \tag{4.7}$$

Combining these, I get an expression of optimal z^* as a function of optimal x^*:

$$\frac{p_z}{p_x} = \frac{1 - L(x^*)}{L_x(x^*)} \cdot \frac{H_z(z^*)}{H(z^*)} \tag{4.8}$$
$$= \frac{x^*}{\delta(x^*)} \cdot \frac{\eta(z^*)}{z^*} \tag{4.9}$$
$$\Longrightarrow \; z^* = x^* \cdot \frac{\eta(z^*)}{\delta(x^*)} \cdot \frac{p_x}{p_z} \tag{4.10}$$

[3]The equations is:

$$\frac{\beta_c^2}{2 \cdot p_z} = \frac{\varepsilon_s}{\sum 1 - L(t_{c,k}^{2016})} \cdot \frac{q_{c,2016}}{p_{2016}} \Longrightarrow \alpha_c = q_{c,2016} - \frac{\beta_c^2}{2 \cdot p_z} \cdot p \cdot \sum \left(1 - L(t_{c,k}^{2016})\right)$$

where η is the elasticity of potential output in z, and δ is the elasticity of (net-of) loss ratio in x[4]. This can be plugged in a FOC to get a necessary conditions for profit maximization:

$$p \cdot L_x(x^*) \cdot H\left(z^*(x^*)\right) = p_x \tag{4.11}$$

This is an implicit function of x^*, given the relative prices and other parameters. To better understand the concept of MCE as a solution for climate challenges, let us differentiate the equation (4.11) w.r.t. the output price p and the optimal MCE input x^*. I get (after some simplification):

$$\frac{\Delta x^*}{\Delta p} = \frac{L_x(x^*) \cdot H\left(z^*(x^*)\right)}{-L_{xx}(x^*) \cdot H\left(z^*(x^*)\right) + L_x(x^*) \cdot H_z(z^*(x^*)) \cdot z_x^*(x^*)} \tag{4.12}$$

Where regularity conditions assure us this ratio is positive (the loss function should be concave in the solution area). Naturally, an increase in output price is related to an increase in the optimal MCE input. However, a significant increase requires a large marginal MCE effect, i.e. $L_x(x^*) \gg 0$ in the numerator. Where this not the case, i.e. $L_x(x^*) \to 0$, increase in price will result in very little increase in x^*. Rather, the grower response would be through changes in z^*.

To specify a loss function with MCE, I assume that each application of kaolin increases the chill count by one portion. Note that the cost of increasing the chill count by one portion depends on the total acreage. The cost of one additional portion per acre is estimated at \$55[5]. In the real world, there is a limit to the potential cooling effects of kaolin clay. Applying more reflective mix on trees already coated with a hefty layer would not be useful. However, as the layers are prone to washing off with winter rain, I take these costs and effects as linear for the model. The total required "extra" chill portions, usually about 15 on a warm year, seems feasible with weekly applications starting early in the winter.

Once the optimal level of MCE –x^*– has been established by solving equation (4.11), a solution for the regular input level –z^*– can be calculated directly (for the algebra details see appendix A.1 and A.2). This results in an implicit supply function for a grower with MCE:

$$q_{c,k}[p] = \left(\alpha_{c,k} + \beta_{c,k} \cdot \frac{-\alpha_{c,k} + \sqrt{\alpha_{c,k}^2 + 2 \cdot \frac{\beta_{c,k}^2}{p_z} \cdot \frac{1-L(x[p])}{L_x(x[p])} \cdot p_x \cdot Acres_{c,k}}}{2 \cdot \beta_{c,k}}\right) \tag{4.13}$$

Note that the $\beta_{c,k}$ terms cancel out, and we are left with an expression to calculate using the coefficients I calibrated before. Then, given the MCE input price p_x and the chill realization for the grower, this expression results in the supplied quantity for any pistachio price.

[4]Note that $L(x)$ is decreasing in x, hence the derivative of the net-of loss function w.r.t. x is positive.

[5]I thank Donald Stewart from UCANR's Agricultural Issues Center for data on material and deployment costs of kaolin clay. Pounds per acre ratios and the expected weekly effect are from (Doll, 2015). I assume a weekly rain event during winter washes off the treatment, which needs to be applied again

Market Demand

In empirical demand estimations, such as the ones I cite below, we usually find either linear or iso-elastic specifications. Linear demand allows for a choke price (i.e. price where zero units are wanted) and demand elasticity that varies with the price, which seems more realistic when modeling large supply disruptions. The demand function in the model is therefore linear:

$$D(p) = a - b \cdot p \tag{4.14}$$

Again, to pin-point a demand function, we need to calibrate these parameters with a point and a slope. For a point, I use the 2016 price and quantity. I assume an elasticity to calculate the slope. Most estimates for pistachio demand elasticity are between -1 and -2. Demand for pistachio is considered elastic, as much of it is exported and it is not a staple food. The elasticity is capped, reflecting relatively low substitutibility because of pistachio's unique flavor. The earliest demand elasticity estimate I found is from the 1970's: Dhaliwal (1972), in Nuckton (1978), estimated it at -1.5. Awondo and Fonsah (2014) try to calculate demand elasticity by using total production and averaging consumption among the US population, using an AIDS based model. They estimate a price elasticity of $-0.96(0.04)$. Gray et al. (2005) cite a report by Lewis, estimating an elasticity ranges of $(-1.66, -1.44)$ for domestic demand, and $(-2.31, -1.59)$ for export demand. Cheng et al. (2017) estimate local demand elasticity using micro-data (the Nielsen barcode data) and get an (uncompensated) price elasticity of $-1.25(0.11)$. Zheng, Saghaian, and Reed (2012) estimate an export demand elasticity of $-1.79(0.34)$, which produces a range quite similar to the 1999 study by Lewis. I chose to combine the latter (more recent) two estimates, given that 2/3 of pistachios are exported. The combined elasticity distribution is $\varepsilon^D \sim N(-1.61, 0.23^2)$. I assume an elasticity of $\varepsilon_D = -1.61$ and later show results with other elasticities as well.

Market Clearing and Welfare Outcomes

Figure 4.1 sketches the short run market model. The linear supply curves take weather as given. On an ideal weather season, the supply curve is S_0. On a year with warm winter, the supply curve is multiplied by a coefficient smaller than one, i.e. shifts left and rotates counter-clockwise, resulting in curve S_1. Without MCE, the intersection of demand with S_1 determines the market equilibrium. Once that is solved, the welfare outcomes-consumer surplus, grower sector profits, and total welfare-are calculated as the areas above or under the appropriate curves.

When MCE technology is available, a modified supply curve starts with a section overlapping S_1, and then "bends" right towards S_0. If demand is high enough, market equilibrium is attained at this bend. Again, the welfare outcomes with MCE are calculated with the equilibrium price and quantity, together with the demand and S_{MCE} curves.

The gains from MCE are the differences between these market outcomes, i.e. the outcomes with MCE minus the outcomes without it. Note that the expansion of supply by

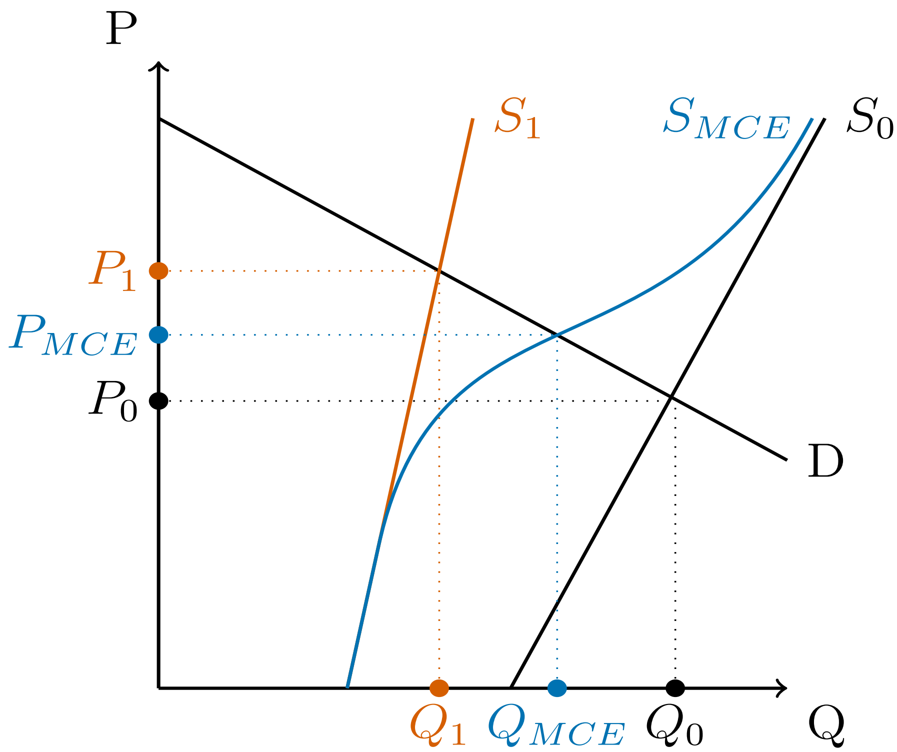

Figure 4.1: Sketch of market model for pistachios with micro-climate Engineering. S_0 is the supply curve under perfect weather. S_1 is the supply curve in a warm year, where yields are lower. S_{MCE} is the supply with micro-climate engineering.

MCE is guaranteed to result in (weakly) positive gains from MCE in terms of total welfare and consumer surplus: the price is lower and quantity is higher. As for the grower sector, it does enjoy extra profits from being able to produce more, but the resulting lower price also decreases its profits from the output that would have been produced anyway without MCE. Therefore, one cannot tell *a priori* if grower profits increase or decrease when MCE is available. The sign and magnitude will need to be determined in the simulations, given the various parameters and functional forms.

The climate prediction data produce a point estimate of chill portions for each year in 2020-2040. For a given set of model parameters and climate predictions for 2020-2040, the model is solved numerically twice for each year in this range. The consumer, grower, and welfare gains are calculated for each year using these two simulations. Using a discount rate of 5%, I can calculate the Net Present Value (NPV) of the MCE gains in 2019. For each scenario, I run this procedure for 100 "independent draws" of 2020-2040 prediction paths. For each one, an entire simulation is run to produce an NPV of the gains. I report the Expected NPV (ENPV), the mean of this distribution, and standard errors around it. More details on the numerical solution of the model can be found in appendix A.3.

Acreage Growth Scenarios

Before I present the simulated welfare gains, there is one more piece in the puzzle. The calibrated model is set with 2016 acreage (329,826 acres). Pistachio acreage through 2020-2040 is likely to be different, and most likely higher than that. However, the model does not include endogenous growth of planted and harvested pistachio acres. To give some bounds on the expected gains, I run the simulations with four different acreage growth scenarios, each specifying a different pistachio acreage growth path until 2040.

All scenarios assume some growth path until 2030, when acreage stabilizes and stays fixed through 2040. The first scenario is "No Growth", meaning that 2020-2040 climate predictions are cast over the 2016 acreage. This should give a lower bound for gains, as acreage is predicted to grow and not shrink. The second scenario is "Low Growth", which sets the yearly growth of harvested acres until the year 2022 at 9.6%, the average rate since 2000, and then sets zero growth (total acreage growth of 75%). The growth until 2022 is attributed to currently planted but not yet bearing acres. This assumes that we are on the brink of a dynamic equilibrium in growth, and therefore no new acres will be planted in California. This scenario should give estimates that are higher than the "No Growth" scenario, but still rather conservative. The third scenario is "High Growth". This one sets the growth rate until 2022 at 14.6%, the average rate since 2010, and then lets pistachio acreage follow the historic path of almonds in California (total acreage growth of 260%). That is, the growth rate of almonds when they had the corresponding pistachio acreage. This very optimistic growth prediction makes the "High Growth" scenario the upper bound for the gains from MCE. One potential concern with acreage growth is that growers might switch new acreage to unaffected counties, or plant more heat tolerant varieties. For this, the "High North" scenario takes the high growth rate, but all new acreage harvested from 2023 is located in an imaginary "North" county, where chill damages are virtually zero. Note that planting in the unaffected north has the same effect on supply as planting a more heat tolerant variety near the existing locations (assuming that the potential output, both in the north and of the new variety, are identical to the current one). This last scenario is, in my opinion, the most plausible in terms of MCE gain magnitudes. A summary of the growth rates is depicted in Figure 4.2. In all scenarios, demand grows by the total rate of acreage growth.

4.4 Simulations Results

I present the Expected NPV of our simulations (and standard deviations) in Table 4.1. The total welfare gains from MCE technologies are positive, for the market as a whole and for consumers specifically. ENPVs of the total welfare gains are between $1.6 billion in the "No Growth" scenario to $4.8 billion in the "High Growth" scenario. Consumer surplus gains range from $2.5 billion to $7.2 million for the same scenarios. The reader might guess by now that the expected gains for growers are negative. Indeed, the increased quantity

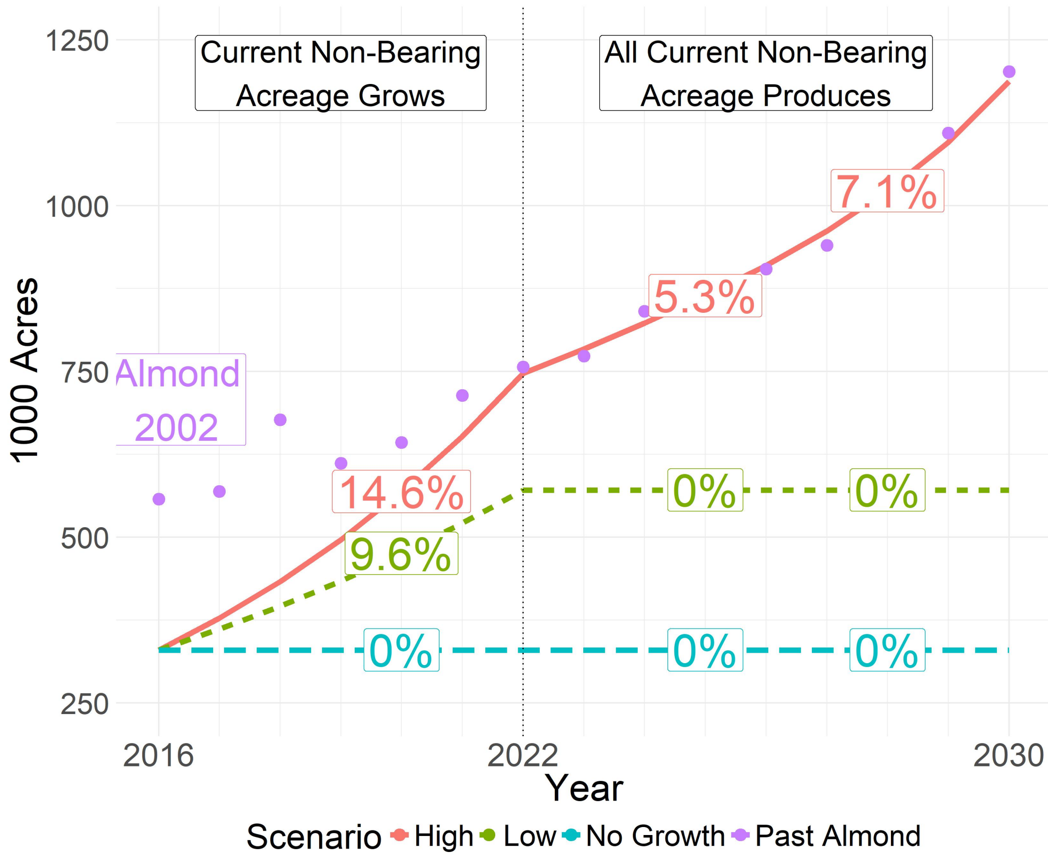

Figure 4.2: Growth rate development for different scenarios.

with MCE is not enough to compensate for the profit loss due to lower price. The gains from MCE for the grower sector are negative, ranging from -0.9 billion dollars in the "No Growth" scenario to -2.4 billion in the "High Growth" scenario. This is true not only for the ENPV calculation, but in general in almost every predicted year and scenario, except for a few with the most extreme adverse weather predictions.

Scenario	Consumer	Grower	Welfare
No Growth	2.5 (0.2)	-0.9 (0.1)	1.6 (0.2)
Low Growth	4.2 (0.4)	-1.5 (0.1)	2.7 (0.3)
High North	5.4 (0.5)	-1.9 (0.1)	3.5 (0.4)
High Same	7.2 (0.7)	-2.4 (0.2)	4.8 (0.6)

Table 4.1: Expected net present value of MCE in billions of US$. Yearly gains in the years 2020-2040 are discounted at 5% yearly and summed to calculate ENPV in 2019. The values presented are the mean (standard deviation) from 100 climate prediction bootstraps.

The average loss for growers is not the result of a distortion. Growers in the model make optimal decisions given the market conditions. MCE expands supply and lowers prices, increasing the total grower revenue. How do we know this? We have a linear demand system where the "perfect weather" equilibrium has elastic demand. On a warm year, the quantity drops to a point with even more elastic demand. The quantity increasing effect of MCE must therefore increase consumer expenditures. Nevertheless, these increased revenues result in lower total profits.

Other elasticity specifications

The main specification assumes a competitive market and takes certain values of elasticities for supply and demand. How would the gains from MCE change under alternative assumptions about market structure and parameters? To test this, I first run the model under the competitive assumptions with different elasticities. I use the values $\varepsilon_S = (0.1, 0.19, 0.3)$ and $\varepsilon_D = (-0.5, -1.1, -1.61, -2)$. Table 4.2 Shows the results for the "High North" scenario in a convenient format. As expected, the more elastic the supply, and the less elastic the demand, consumer gains (and total welfare gains) from MCE increase, as do grower profits losses. While the profits and surplus vary a lot between the different elasticity pairs, the movement is opposite in such way that the total expected welfare gains are relatively stable.

Grower ; Consumer Welfare	$\varepsilon_D =$ -2.0	$\varepsilon_D =$ -1.61	$\varepsilon_D =$ -1.1	$\varepsilon_D =$ -0.5
$\varepsilon_S = 0.10$	-0.7 ; 3.7 3.0	-1.4 ; 4.5 3.1	-3.2 ; 6.4 3.3	-9.2 ; 13.1 3.9
$\varepsilon_S = 0.19$	-1.0 ; 4.4 3.4	**-1.9 ; 5.4** **3.5**	-3.8 ; 7.5 3.7	-9.7 ; 13.9 4.2
$\varepsilon_S = 0.30$	-1.9 ; 6.1 4.2	-3.0 ; 7.3 4.3	-5.3 ; 9.7 4.5	-11.2 ; 16.2 4.9

Table 4.2: Expected NPV gains (billion $US) from MCE under varying elasticities in the "High North" scenario. Top left is grower gains, top right is consumer gains, bottom is total welfare gains. The emphasized numbers correspond to the main specification.

Introducing Market Power

So far, I assumed that consumers buy directly from growers, and the market is competitive. In fact, pistachios are processed and marketed by intermediaries, and it has been reported that about half of the output is marketed by one firm (Blank, 2016). Combined with high entry costs (no income for at least 6 years as trees mature), it would seem plausible that some market power is being exercised. Therefore, it is interesting to examine the gains from MCE under some degree of market power in the supply chain. The purpose of this exercise is not to try and evaluate the existing market power in pistachios, or to assess the potential welfare effects of market power relative to a competitive market. Rather, the question is: what would be the gains from MCE if market power exists?

To include market power in the model, I use a flexible framework with an intermediary or middleman which can have market power on consumers (see Just, Hueth, and Schmitz 2005, p. 386-388 and Sexton and Zhang 2001). In fact, the model can also accommodate market power on the growers (monopsonistic power, e.g. from large retail chains). For simplicity, and since determining a range for the real degree of monopsonistic market power is complicated, I only use the monopolistic market power part. This intermediary manipulates the price for growers and consumers to maximize its profit. Maximum profit is attained when the intermediary equates the marginal revenue from sales to consumers with marginal outlay paid to growers plus extra costs in the supply chain. The result is a fixed ratio between the price for consumers and the marginal cost of the intermediary, which is the grower price plus the processing and handling costs. This is, of course, an extension of the celebrated work by Lerner (1934), who realized that the price-cost margin is evidence of monopoly strength, and that this margin should – in theory – be equal to the inverse of demand elasticity. The explicit derivation, relating the marginal revenue to price and elasticity, is a well known textbook result (e.g. Carlton and Perloff, 2005, p. 92). The equation linking grower and consumer prices is:

$$p^{CONSUMER} = \left(p^{GROWER} + \delta\right) \times \left(1 + \frac{\psi}{\varepsilon^D}\right)^{-1} \tag{4.15}$$

where $\psi \in [0,1]$ is a market power measure w.r.t. the consumer sector (oligopolistic market power), where zero is no market power and 1 is monopoly. ε^D is the price elasticity of demand, which is negative, making the term in parenthesis smaller than one[6]. δ are added costs in the supply chain from grower to consumer, and are assumed to be fixed.

The actual measure of market power ψ for pistachios is unknown. However, it is useful to think of a reasonable upper bound for it. I run the simulations with $\psi = 0.5$ as upper bound, and $\psi = 0.25$ for a middle point between the upper bound and the competitive market simulations reported above. Applying market power means limiting the supplied quantities. This might increase or decrease the "raw" grower profits, while creating positive profits for the intermediary. To get a sense of the total oligopsonist gains, I add both the grower and intermediary gains together, resulting in "Agribusiness" gains. Table 4.3 shows the results from these simulations, ordered by the gains for agribusiness sector. It turns out that only under the rather extreme ends of our parametric range, including market power measure, does the agribusiness sector see positive gains from MCE.

To get a sense of the effect of all variables on the gains from MCE, Figure 4.3 shows the "ceteris paribus" picture for gains under the "High Same" scenario. This plot shows the outcomes from running the model on 200 variable parameter combinations. Note that the

[6]This number is greater than zero since I am assuming elastic demand. Accordingly, I do not run simulations with market power for $\varepsilon_D = -0.5$

ε_D	ε_S	ψ	Agribusiness	Consumer	Welfare
-2.00	0.10	0.50	3.2	5.7	8.9
-2.00	0.19	0.50	2.2	8.3	10.5
-1.61	0.10	0.50	2.2	8.6	10.8
-2.00	0.10	0.25	1.0	4.6	5.6
-2.00	0.19	0.25	0.5	5.9	6.4
-1.61	0.10	0.25	0.1	6.1	6.3
-2.00	0.30	0.50	-0.1	13.6	13.4
-2.00	0.10	0.00	-0.2	3.7	3.6
-1.61	0.19	0.50	-0.2	12.8	12.6
-2.00	0.19	0.00	-0.5	4.4	4.0
-2.00	0.30	0.25	-0.8	8.7	8.0
-1.61	0.19	0.25	-0.8	7.9	7.1
-1.61	0.10	0.00	-0.9	4.6	3.7
-2.00	0.30	0.00	-1.2	6.0	4.8
-1.61	**0.19**	**0.00**	**-1.3**	**5.4**	**4.1**
-1.61	0.30	0.00	-2.3	7.2	4.9
-1.10	0.10	0.00	-2.6	6.6	4.0
-1.61	0.30	0.25	-2.8	11.6	8.7
-1.10	0.10	0.25	-3.1	10.9	7.8
-1.10	0.19	0.00	-3.2	7.6	4.4
-1.10	0.30	0.00	-4.5	9.7	5.2
-1.61	0.30	0.50	-4.8	20.3	15.4
-1.10	0.19	0.25	-5.4	14.0	8.6
-1.10	0.10	0.50	-6.3	21.8	15.5
-1.10	0.30	0.25	-9.4	19.4	10.1
-1.10	0.19	0.50	-16.9	32.7	15.8
-1.10	0.30	0.50	-30.6	46.7	16.0

Table 4.3: Expected gains (billion $US) from MCE under varying elasticities and market power in the "High North" scenario. Agribusiness gains are grower plus intermediary gains. The emphasized numbers correspond to the main specification.

range of drawn parameters is wider than our specified parameter ranges, which I think are reasonable for the pistachio market. This is why we see many results with positive grower gains from MCE.

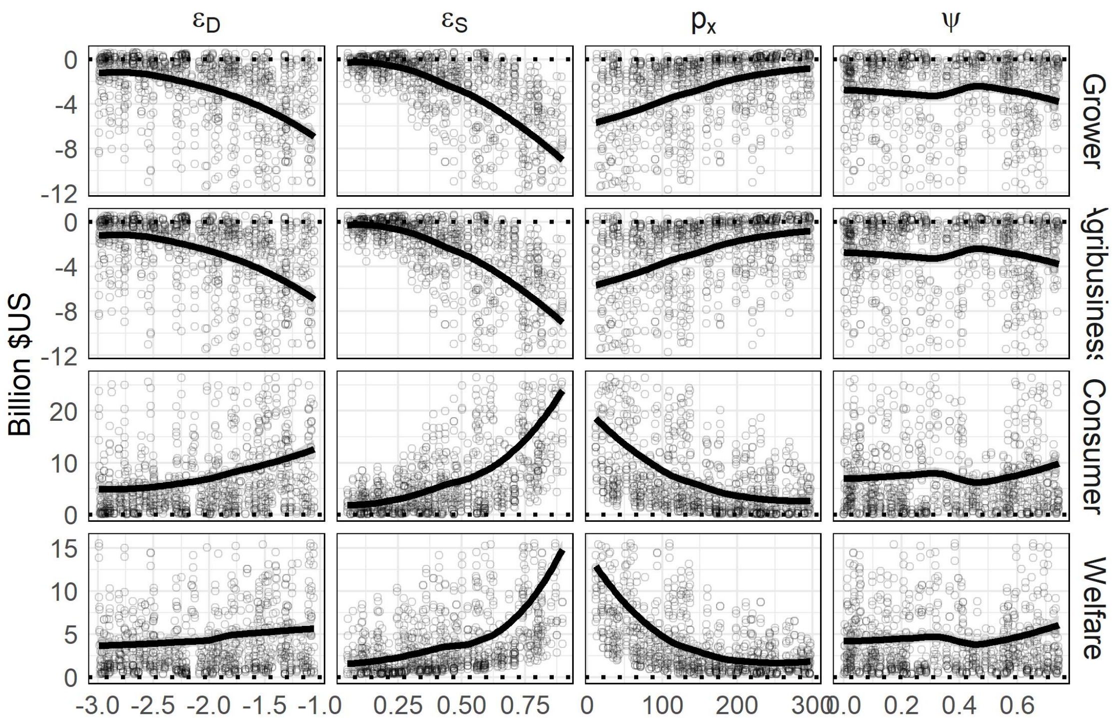

Figure 4.3: Effect of model variables on MCE gains. Each point is an ENPV of a model run. Each of the 200 randomly drawn parameter combinations are run on 5 random climate bootstraps. The result are 1000 ENPV values shown here.

The solid line in each panel is the prediction from a local polynomial regression, a sort of empirical derivative. Grower and Agribusiness profit losses from MCE are greater when demand is more elastic and supply is less elastic, as expected. Consumer surplus gains behave in the opposite way. The total welfare gains from MCE overall seem pretty stable with demand elasticity in the specified ranges, but increases when supply is more elastic. The price of MCE, p_x, ranges from 10 to 300 in the plot. As it increases, consumer and total welfare gains from MCE decrease, as less MCE input is used by growers, as expected. The profit and agribusiness gains increase (while still negative) and then plateau towards zero, probably as the price increases lower the total use of MCE to very low levels.

4.5 Discussion and Conclusion

MCE could help overcome a climate challenge for California pistachios. I model the market and assess the potential welfare gains from a reflective coating technology that lowers the effective temperatures in pistachio orchards. The expected NPV in 2019, for the gains from this technology between 2020 and 2040, is predicted to be around $2.7-3.5 billion. These come from consumer surplus gains, as the total gains for growers in the main specifications are negative. The latter result is not unheard of in agricultural settings, where a negative supply shock can actually increase grower profits. For example, Carter et al. (1981) show that the 1979 labor strikes in California actually increased revenues and profits for lettuce growers. The simulation results shows the flip side of the coin: solving a (weather generated) supply shock can lower grower profits.

While less tangible (and taxable) than actual registered profits, consumer surplus gains are real economic gains enjoyed by the public. This point holds even when discussing a narrower welfare framework for California alone. Part of the modeled gains in consumer surplus are enjoyed elsewhere, as the majority of pistachio output is currently exported. However, export demand is usually considered more elastic than domestic demand, making the share of local consumer surplus gains disproportionate to the share of local consumption. At a share of 1/3 of total consumption, let us assume that Californians still enjoy half of the consumer surplus gains from MCE (and the entire grower gains). Adjusting Table 4.1, the total welfare gains in California are strictly negative when the demand is unrealistically inelastic, $\varepsilon_D = -0.5$, and strictly positive for more realistic demand assumptions ($\varepsilon_D < -1.1$).

The scope of consumer surplus gains brings us to the potential gains from public investment in R&D for MCE solutions. With social returns from investments largely exceeding private ones, this type of research is a good candidate for prioritizing in public research fund allocation (Alston, Norton, and Pardey, 1995, p. 491). The case for public research is made stronger by the fact that there seems to be little private incentive to invest in MCE, at least in this case. I see MCE technologies mostly as an adaptation of existing ones to solve a climate problem. Therefore, innovations in the field would be hard to make proprietary by the innovator. Moreover, innovators are likely to come from the industry: a large growing firm would have the resources and access to enough pistachio acreage to run experiments and develop new MCE solutions. But if this firm sees that a world with MCE (adopted by everyone) is worse, why invest in innovation? Adding market power to the equation makes an even stronger potential case for public R&D: the total welfare gains are higher, and the incentives for innovation could be even lower.

What might be the implications of MCE technologies in a broader sense? One could imagine, with further agronomic research, other MCE technologies applied to other fruit and nut crops, and even for annuals such as corn or soybeans. Of course, these are less profitable than pistachios, but they face similar challenges, and MCE solutions are not necessarily very expensive. Other implications could be with the distribution of climate change damage incidence. Technologies might only be available (and affordable) to growers in countries better off financially, further exacerbating international income disparities. An interesting potential for MCE technologies could be in accelerating the transition of agricultural practices closer to the poles, sometimes referred to as the "crop migration" (Zilberman et al., 2004). For example, MCE solutions for frost could accelerate the expansion of viticulture to higher latitudes.

The simulation based valuation methodology in this chapter has its caveats. Modeling supply and demand as linear is obviously a simplification. The assumptions on growth and distribution of acreage are based on past growth patterns, and might not reflect unexpected future changes in market conditions. The future chill predictions are in line with other predictions by climatologists, yet might fail to materialize. Nevertheless, by choosing various scenarios, basing the parameter ranges in the literature, and choosing conservatively when possible, I believe to have gotten a reasonable range for the potential gains from MCE in California pistachios. They are in the low billions for a crop of secondary importance to California agriculture. I believe this shows a great potential of MCE technologies for climate change adaptation in general.

References

Alston, J.M., G.W. Norton, and P.G. Pardey. 1995. *Science under scarcity: principles and practice for agricultural research evaluation and priority setting*. Cornell University Press.

Alston, J.M., P.G. Pardey, J.S. James, and M.A. Andersen. 2009. "The Economics of Agricultural R&D." *Annual Review of Resource Economics* 1:537–566.

Anaman, K.A., and S.C. Lellyett. 1996. "Assessment of the benefits of an enhanced weather information service for the cotton industry in Australia." *Meteorological Applications* 3:127–135.

Auffhammer, M., and W. Schlenker. 2014. "Empirical studies on agricultural impacts and adaptation." *Energy Economics* 46:555 – 561.

Awondo, S., and G. Fonsah. 2014. "Short Run and Long Run dynamics in the Demand of US Tree Nuts." In *2014 Annual Meeting, February 1-4, 2014, Dallas, Texas*. Southern Agricultural Economics Association, 162465.

Ayars, J., A. Fulton, and B. Taylor. 2015. "Subsurface drip irrigation in California—Here to stay?" *Agricultural Water Management* 157:39–47.

Babcock, B.A. 1990. "The Value of Weather Information in Market Equilibrium." *American Journal of Agricultural Economics* 72:63–72.

Baldocchi, D., and E. Waller. 2014. "Winter fog is decreasing in the fruit growing region of the Central Valley of California." *Geophysical Research Letters* 41:3251–3256.

Baldocchi, D., and S. Wong. 2008. "Accumulated winter chill is decreasing in the fruit growing regions of California." *Climatic Change* 87:153–166.

Barrows, G., S. Sexton, and D. Zilberman. 2014. "The impact of agricultural biotechnology on supply and land-use." *Environment and Development Economics* 19:676–703.

Beede, B., and D. Doll. 2016. "Sun Reflecting Products for Increased Winter Chill?" Working paper, University of California Cooperative Extension.

Benmoussa, H., E. Luedeling, M. Ghrab, J.B. Yahmed, and M.B. Mimoun. 2017. "Performance of pistachio (Pistacia vera L.) in warming Mediterranean orchards." *Environmental and experimental botany* 140:76–85.

Binford, M.W., A.L. Kolata, M. Brenner, J.W. Janusek, M.T. Seddon, M. Abbott, and J.H. Curtis. 1997. "Climate variation and the rise and fall of an Andean civilization." *Quaternary research* 47:235–248.

Blank, S.C. 2016. "The Economic Outlook for California Pistachios." In L. Ferguson and D. R. Haviland, eds. *The Pistachio Production Manual*. Oakland, California: University of California, Division of Agriculture and Natural Resources, chap. 1, pp. 3–10.

Boryan, C.G., Z. Yang, and P. Willis. 2014. "US geospatial crop frequency data layers." In *Agro-geoinformatics (Agro-geoinformatics 2014), Third International Conference on*. IEEE, pp. 1–5.

Carleton, T.A., and S.M. Hsiang. 2016. "Social and economic impacts of climate." *Science* 353:aad9837.

Carlton, D.W., and J.M. Perloff. 2005. *Modern industrial organization*. Pearson Higher Ed.

Carter, C.A., D.L. Hueth, J.W. Mamer, and A. Schmitz. 1981. "Labor strikes and the price of lettuce." *Western Journal of Agricultural Economics*, pp. 1–14.

CDFA. 2017. *California Agricultural Statistics Review, 2016-2017*.

CEDA. 2016. Downloaded Data, retrived October 20th, 2016.

Chambers, R.G., and E. Lichtenberg. 1994. "Simple econometrics of pesticide productivity." *American Journal of Agricultural Economics* 76:407–417.

Cheng, G., S. Dharmasena, O. Capps Jr, et al. 2017. "The Taste for Variety: Demand Analysis for Nut Products in the United States." In *2017 Annual Meeting, February 4-7, 2017, Mobile, Alabama*. Southern Agricultural Economics Association, 252719.

CIMIS. 2018. Downloaded Data, retrived July 20th, 2018.

Cirac-Claveras, G. 2019. "Weather Satellites: Public, Private and Data Sharing. The Case of Radio Occultation Data." *Space Policy* 47:94–106.

Cohen-Vogel, D.R., D.E. Osgood, D.D. Parker, and D. Zilberman. 1998. "The California Irrigation Management Information System (CIMIS): Intended and Unanticipated Impacts of Public Investment." *Choices* 13.

CPRB. 2009. "Guidelines for California Pistachio Growers." Working paper, California Pistachio Research Board.

Cullis, C., and K.J. Kunert. 2017. "Unlocking the potential of orphan legumes." *Journal of experimental botany* 68:1895–1903.

Dhaliwal, H. 1972. "An Econometric Investigation of Demand Interrelationships Among Tree Nuts and Peanuts." PhD dissertation, Oregon State University.

Doll, D. 2015. "Kaolin Clay May be Useful in Increasing Chill Accumulation in Pistachios." Working paper, University of California Cooperative Extension.

—. 2017. "Regional Considerations for Pistachio Production." In *Advances in Pistachio Production Short Course*. University of California Division of Agriculture and Natural Resources, Online; Retreived June 19th, 2018 from http://ucanr.edu/sites/PistachioShortCourse/files/274446.pdf.

Elloumi, O., M. Ghrab, H. Kessentini, and M.B. Mimoun. 2013. "Chilling accumulation effects on performance of pistachio trees cv. Mateur in dry and warm area climate." *Scientia Horticulturae* 159:80–87.

Elzhov, T.V., K.M. Mullen, A.N. Spiess, and B. Bolker. 2016. *minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds*. R package version 1.2-1.

Erez, A. 2000. "Bud dormancy; phenomenon, problems and solutions in the tropics and subtropics." In *Temperate fruit crops in warm climates*. pp. 17–48.

Erez, A., and S. Fishman. 1997. "Dynamic Model Chilling Portions." Retrieved June 16h, 2016. http://ucanr.edu/sites/fruittree/files/49319.xls.

Erez, A., S. Fishman, G. Linsley-Noakes, and P. Allan. 1989. "The dynamic model for rest completion in peach buds." In *II International Symposium on Computer Modelling in Fruit Research and Orchard Management 276*. pp. 165–174.

Fishman, S., A. Erez, and G. Couvillon. 1987a. "The temperature dependence of dormancy breaking in plants: computer simulation of processes studied under controlled temperatures." *Journal of Theoretical biology* 126:309–321.

—. 1987b. "The temperature dependence of dormancy breaking in plants: mathematical analysis of a two-step model involving a cooperative transition." *Journal of Theoretical Biology* 124:473–483.

Glozer, K. 2016. "The Dynamic Model and Chill Accumulation." Retrieved June 16h, 2016.

Gordon, B., O. Sambucci, I. Trilnick, and D. Zilberman. 2018. "Research, Innovation, Supply Chains, and Precision Agriculture in California." In P. L. Martin, R. E. Goodhue, and B. D. Wright, eds. *California Agriculture: Dimensions and Issues*. Giannini Foundation of Agricultural Economics and University of California Division of Agriculture and Natural Resources, chap. 17.

Gray, R.S., D.A. Sumner, J.M. Alston, H. Brunke, A.K. Acquaye, et al. 2005. "Economic consequences of mandated grading and food safety assurance: ex ante analysis of the federal marketing order for California pistachios." *Giannini Foundation Monograph* 46.

Hu, Y., E.A. Asante, Y. Lu, A. Mahmood, N.A. Buttar, and S. Yuan. 2018. "Review of air disturbance technology for plant frost protection." *International Journal of Agricultural and Biological Engineering* 11:21–28.

IPCC. 2013. *Summary for policymakers*, Cambridge, UK & New York, NY, USA: Cambridge University Press.

Irvine, P.J., B. Kravitz, M.G. Lawrence, and H. Muri. 2016. "An overview of the Earth system science of solar geoengineering." *Wiley Interdisciplinary Reviews: Climate Change* 7:815–833.

Jarvis-Shean, K. 2017. "How Potential Changes in Climate Could Affect Pistachio Production." In *Advances in Pistachio Production Short Course*. University of California Division of Agriculture and Natural Resources, Online; Retreived June 19th, 2018 from http://ucanr.edu/sites/PistachioShortCourse/files/274446.pdf.

Johnson, S.R., and M.T. Holt. 1997. "The value of weather information." In R. W. Katz and A. H. Murphy, eds. *Economic Value of Weather and Climate Forecasts*. Cambridge University Press, pp. 75–108.

Just, R.E., D.L. Hueth, and A. Schmitz. 2005. *The welfare economics of public policy: a practical approach to project and policy evaluation*. Edward Elgar Publishing.

Kallsen, C.E. 2017. "Temperature-related Variables Associated with Yield of 'Kerman' Pistachio in the San Joaquin Valley of California." *HortScience* 52:598–605.

Kingsbury, N. 2009. *Hybrid: the history and science of plant breeding*. University of Chicago Press.

Klockow, K.E., R.A. McPherson, and D.S. Sutter. 2010. "On the economic nature of crop production decisions using the Oklahoma Mesonet." *Weather, Climate, and Society* 2:224–236.

Kolata, A.L., and C. Ortloff. 1989. "Thermal analysis of Tiwanaku raised field systems in the Lake Titicaca Basin of Bolivia." *Journal of Archaeological Science* 16:233–263.

Lave, L.B. 1963. "The value of better weather information to the raisin industry." *Econometrica* 31:151–164.

Leard, B., and K. Roth. 2019. "Voluntary Exposure Benefits and the Costs of Climate Change." *Journal of the Association of Environmental and Resource Economists* 6:151–185.

Lerner, A.P. 1934. "The Concept of Monopoly and the Measurement of Monopoly Power." *The Review of Economic Studies* 1:157–175.

Lichtenberg, E., and D. Zilberman. 1986. "The econometrics of damage control: why specification matters." *American Journal of Agricultural Economics* 68:261–273.

Lu, Y., Y. Hu, C. Zhao, and R.L. Snyder. 2018. "Modification of Water Application Rates and Intermittent Control for Sprinkler Frost Protection." *Transactions of the ASABE* 61:1277–1285.

Luedeling, E. 2017. *chillR: Statistical Methods for Phenology Analysis in Temperate Fruit Trees*. https://CRAN.R-project.org/package=chillR.

—. 2012. "Climate change impacts on winter chill for temperate fruit and nut production: a review." *Scientia Horticulturae* 144:218–229.

Luedeling, E., M. Zhang, and E.H. Girvetz. 2009. "Climatic Changes Lead to Declining Winter Chill for Fruit and Nut Trees in California during 1950–2099." *PLOS ONE* 4:1–9.

MacAuley, M.K. 2005. "The Value of Information: A Background Paper on Measuring the Contribution of Space-Derived Earth Science Data to National Resource Management." Discussion paper 05-26, Resources For the Future.

Norton, G.W., and J.S. Davis. 1981. "Evaluating returns to agricultural research: a review." *American Journal of Agricultural Economics* 63:685–699.

Nuckton, C.F. 1978. "Demand Relationships for California Tree Fruits, Grapes, and Nuts: A Review of Past Studies." Working paper, University of Califronia, Davis.

Olen, B., J. Wu, and C. Langpap. 2015. "Irrigation decisions for major west coast crops: Water scarcity and climatic determinants." *American Journal of Agricultural Economics* 98:254–275.

Olmstead, A.L., and P.W. Rhode. 2011. "Adapting North American wheat production to climatic challenges, 1839–2009." *Proceedings of the National Academy of sciences* 108:480–485.

Parker, D., D. Cohen-Vogel, D. Osgood, D. Zilberman, et al. 2000. "Publicly funded weather database benefits users statewide." *California Agriculture* 54:21–25.

Parker, D., D. Zilberman, D. Cohen-Vogel, and D. Osgood. 1996. "The Economic Costs and Benefits Associated with the California Irrigation Management Information System (CIMIS)." Working paper, Report to the California Department of Water Resources.

Parker, D.D., and D. Zilberman. 1993. "Hedonic estimation of quality factors affecting the farm-retail margin." *American Journal of Agricultural Economics* 75:458–466.

Persson, U.M. 2015. "The impact of biofuel demand on agricultural commodity prices: a systematic review." *Wiley Interdisciplinary Reviews: Energy and Environment* 4:410–428.

Pope, K. 2015. "Fruit & Nut Crop Chill Portions Requirements." Working paper, University of California Division of Agriculture and Natural Resources.

Pope, K.S., V. Dose, D. Da Silva, P.H. Brown, and T.M. DeJong. 2015. "Nut crop yield records show that budbreak-based chilling requirements may not reflect yield decline chill thresholds." *International Journal of Biometeorology* 59:707–715.

Russo, C., R.D. Green, and R.E. Howitt. 2008. "Estimation of Supply and Demand Elasticities of California Commodities." Working paper, UC Davis: Department of Agricultural and Resource Economics.

Schlenker, W., and M.J. Roberts. 2009. "Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change." *Proceedings of the National Academy of Sciences* 106:15594–15598.

Sexton, R.J., and M. Zhang. 2001. "An assessment of the impact of food industry market power on US consumers." *Agribusiness* 17:59–79.

Sexton, S.E., Z. Lei, D. Zilberman, et al. 2007. "The economics of pesticides and pest control." *International Review of Environmental and Resource Economics* 1:271–326.

Sharma, R., S.V.R. Reddy, and S. Datta. 2015. "Particle films and their applications in horticultural crops." *Applied Clay Science* 116:54–68.

Stewart, T.R., R.W. Katz, and A.H. Murphy. 1984. "Value of weather information: A descriptive study of the fruit-frost problem." *Bulletin of the American Meteorological Society* 65:126–137.

Taylor, R., D. Parker, and D. Zilberman. 2014. "Contribution of University of California Cooperative Extension to drip irrigation." *ARE Update* 18:5–8.

Thompson, J.F., and K. Adel A. 2016. "Harvesting, Transporting, Procesing, and Grading." In L. Ferguson and D. R. Haviland, eds. *The Pistachio Production Manual*. Oakland, California: University of California, Division of Agriculture and Natural Resources, chap. 17, pp. 189–195.

Tribbia, J.J. 1997. "Weather prediction." In R. W. Katz and A. H. Murphy, eds. *Economic Value of Weather and Climate Forecasts*. Cambridge University Press, pp. 1–18.

Trilnick, I. 2018. *miniChill: calculating chill portions with C++*. R package version 1.0.

U.S. Department of Agriculture. 2013. "Farm and Ranch Irrigation Survey (2013)." *Volume 3* AC-12-SS-1.

Vekshin, A. 2014. "California water prices soar for farmers as drought grows." *Bloomberg News* retrieved 7/2/2019 from `www.bloomberg.com/news/articles/2014-07-24/california-water-prices-soar-for-farmers-as-drought-grows`.

Waterfield, G., and D. Zilberman. 2012. "Pest management in food systems: an economic perspective." *Annual Review of Environment and Resources* 37:223–245.

Zhang, J., and C. Taylor. 2011. "The dynamic model provides the best description of the chill process on 'Sirora'pistachio trees in Australia." *HortScience* 46:420–425.

Zheng, Z., S. Saghaian, and M. Reed. 2012. "Factors affecting the export demand for US pistachios." *International Food and Agribusiness Management Review* 15:139–154.

Zilberman, D., X. Liu, D. Roland-Holst, and D. Sunding. 2004. "The economics of climate change in agriculture." *Mitigation and Adaptation Strategies for Global Change* 9:365–382.

Zilberman, D., A. Schmitz, G. Casterline, E. Lichtenberg, and J.B. Sievert. 1991. "The Economics of Pesticide Use and Regulation." *Science* 253:518–522.

Appendix A

MCE Model Details

A.1 Getting an expression for z^*

The representative county grower maximizes profit. γ_i represent acreage growth for the grower by the simulated year, and needs to pre-multiply p_x and the production coefficients.

$$\max_{(x\,,\,z)} \pi = \gamma_i \cdot p \cdot [1 - L(x_i)] \cdot (\alpha + \beta \cdot \sqrt{z_i}) - p_z^T \cdot z_i - \gamma_i \cdot p_x \cdot x$$

Taking first order conditions:

$$\begin{aligned} \gamma_i \cdot \ p \cdot L_x(x) \cdot (\alpha + \beta \cdot \sqrt{z_i}) &= \gamma_i \cdot p_x \\ \gamma_i \cdot \ p \cdot (1 - L(x)) \cdot \frac{\beta}{2 \cdot \sqrt{z_i}} &= p_z \end{aligned}$$

Combining them:

$$\begin{aligned} \frac{p_z}{\gamma_i \cdot p_x} &= \frac{\gamma_i \cdot \ (1 - L(x))}{\gamma_i \cdot \ L_x(x)} \cdot \frac{\frac{\beta}{2 \cdot \sqrt{z_i}}}{\alpha + \beta \cdot \sqrt{z_i}} \\ \implies \alpha + \beta \cdot \sqrt{z_i} &= \frac{p_x}{p_z} \cdot \frac{(1 - L(x))}{L_x(x)} \cdot \frac{\beta}{2 \cdot \sqrt{z_i}} \cdot \gamma_i \\ \alpha \cdot \sqrt{z_i} + \beta \cdot z_i &= \frac{p_x}{p_z} \cdot \frac{(1 - L(x))}{L_x(x)} \cdot \frac{\beta}{2} \cdot \gamma_i \end{aligned}$$

$$\beta \cdot z_i + \alpha \cdot \sqrt{z_i} - \frac{p_x}{p_z} \cdot \frac{(1 - L(x))}{L_x(x)} \cdot \frac{\beta \cdot \gamma_i}{2} = 0$$

$$\sqrt{z_i^*} = \frac{-\alpha \pm \sqrt{\alpha^2 + 2 \cdot \frac{\beta^2}{p_z} \cdot \frac{1 - L(x)}{L_x(x)} \cdot p_x \cdot \gamma_i}}{2 \cdot \beta}$$

Note that, later one, the β in the denominator cancels out in the production function. Thus we are not required to calculate it directly. Since $\alpha > 0$, and z_i has to be a real number, I only consider the positive solution.

A.2 One Optimal Solution For Grower With MCE

A few challenges arise with our problem specification. First, since I am adding "artificial" chill portions (x) to a natural chill realization, the optimal level of MCE could turn out negative. In this case, the grower is set to supply the quantity with zero MCE (the no-MCE case) as chill portions can be "bought" but not sold. A second challenge is that, given that the production function is not quasi-concave over its support, there might not be an internal solution at all. In this case, I show below that this must be due to the price of input p_x being too high to justify any level of x, and therefore the grower is again set to supply the no-MCE quantity. In Figure 4.1, this is the area where the supply with MCE coincides with the no-MCE supply. A third challenge is that, for the same non-regularity, there might be more than one solution for x^* which solves the grower FOC. Visually, this is evident in Figure 4.1, where the slopes of the MCE supply curve at different prices are not necessarily unique. In this case, I prove below that there are up to two solutions for the FOC, and the highest one is a local maximum and profit maximizing one. In the numerical solution, I make sure to choose the higher one when there are two.

Proposition 1. *There is a unique choice of MCE level for a glower which maximizes his profits.*

Lemma 1. *The value of marginal productivity of the MCE input x is positive and has a maximum.*

Proof. The VMPx is: $p \cdot L_x(x^*) \cdot H\left(z^*(x^*)\right)$ All the components are positive and continuous, therefore VMPx is positive and continuous. Note that $\lim_{x\to\infty} VMPx = 0$ and $\lim_{x\to-\infty} VMPx = 0$. Therefore, there exists some closed interval of x for which at least some values of VMPx are weakly greater than any value outside that interval. On that closed interval, the continuous VMPx attains a maximum value by the extreme value theorem. Since there are values inside the interval that are weakly greater than any values outside of it, that maximum value is the maximum of VMPx for any x. □

As a corollary from this lemma, I can also say that no solution for the equation $p \cdot L_x(x^*) \cdot H\left(z^*(x^*)\right) = p_x$ means that the price p_x (which is positive) is just too high for any choice of x: $VMPx < p_x \; \forall x$. In this case, the profit maximizing solution would be to use zero x.

Lemma 2. *VMPx is unimodal, i.e. has one local maximum.*

Proof. From the previous lemma, we know VMPx attains a maximum and has at least one critical point. To find how many critical points there could be, derive the VMPx and equate to zero, setting $t = exp(m + nx)$:

$$\frac{n^2 \cdot t}{(1+t)^4} \cdot (1-t) \cdot \left[\alpha + \sqrt{\alpha^2 + 2 \cdot \frac{\beta^2}{p_z} \cdot \frac{1+t}{n} \cdot p_x \cdot \gamma_i}\right] + \frac{n \cdot t}{(1+t)^2} \cdot \frac{2 \cdot \frac{\beta^2}{p_z} \cdot \frac{n}{n} \cdot p_x \cdot t}{2\sqrt{\alpha^2 + 2 \cdot \frac{\beta^2}{p_z} \cdot \frac{1+t}{n} \cdot p_x}} = 0$$

$$\frac{n}{(1+t)^2} \cdot (1-t) \cdot \left[\alpha + \sqrt{\alpha^2 + 2 \cdot \frac{\beta^2}{p_z} \cdot \frac{1+t}{n} \cdot p_x \cdot \gamma_i}\right] + \frac{\frac{\beta^2}{p_z} \cdot p_x \cdot t}{\sqrt{\alpha^2 + 2 \cdot \frac{\beta^2}{p_z} \cdot \frac{1+t}{n} \cdot p_x}} = 0$$

Note that the right side is always positive, and the left side is also positive with the exception of $1-t$. Morover, $1-t$ is monotonically decreasing in x. Therefore, there will be only one point where this derivative is zero. Hence VMP_x has only one critical point, which must be a maximum according to the previous lemma. Note that this maximum is attained when $t > 1$, i.e. at an x value which is higher than the location parameter of the logistic distribution which is taken as the damage function. This is the range where the logistic distribution is in fact concave. □

Lemma 3. *The grower FOC has up to two solutions.*

Proof. Since VMPx is unimodal, the FOC: $p \cdot L_x(x^*) \cdot H(z^*(x^*)) = p_x$ has up to two solutions. □

Lemma 4. *If there are two internal solutions, the one with higher x is more profitable for the grower.*

Proof. Two internal solutions means that p_x intersects the unimodal $VMPx$ at two points. These intersections create an interval between the intersection points, where $VMPx > p_x$ for any x in the interval. Moving from the lower to the higher, the grower earns the difference between value of marginal productivity of x and it's price. Hence the intersection with higher x has higher profits. Also note that z^* is the same in both cases, so expenditures on z do not change. □

In the code, I make sure to verify that the higher solution for the FOC of each county-quintile is chosen if more than one exists. The numerical solution of the entire equation system, which includes a market clearing equation, seems to always reach the higher root anyway. This might be because the lower solution is likely to be in an area where the damage function is actually convex in x, and the numerical solver looks for a steady state.

Proof of proposition 1. If there is no internal solution for the FOC, or if the solution is negative (not feasible), there is one optimal solution: $x^* = 0$. If there is a positive internal solution for the FOC, lemmas 1 - 4 assure us that there is only one level of x^* where the grower maximizes profits. □

A.3 Numerical Solution

I model the entire supply with 30 growers: five for each of the six counties. Each of these five represents a county chill quintile realization. The total market supply is the sum of these supplies. When simulating without MCE, the linear supply functions can be added directly, and only a market clearing equation needs to be solved.

Simulating the model with MCE is more complicated, as the implicit supply functions of our 30 representative growers are not additive. I have 30 equations such as equation (4.11) to determine the equilibrium quantity of MCE input x^*_{cd} for each county-quintile. These values are then used to calculate the county-quintile supplied quantities, such as in equation (4.13). The sum of this quantity is equated with demand to clear for a price. This system of 31 equations is numerically solved for one price and 30 levels of x^*_{cd}, which translate to supplied quantities. A solution for this system is the market equilibrium. The consumer surplus is calculated, as before, using the area under the linear supply curve. For grower profits, I need the area under a supply curve, or sum of areas under the 30 supply curves. However, these supply functions are implicit, and I cannot directly integrate them. I approximate this integral by solving for each grower's output for a range of 20 equally distanced prices from zero to the equilibrium price. I then create rectangles using these points, and sum them to approximate for the grower profits.

The model is run for each year in 2020-2040, and net present value is calculated. This procedure produces one simulation result for each set of pre-defined parameters. However, the yearly predictions I use are not intended to forecast the weather in specific years (e.g. predicting the chill in 2035), but rather present the climate trend and variation around it. The climate predictions are therefore a stochastic input, making the otherwise deterministic market model and simulations stochastic as well. We are interested in the expected gains from MCE, given the predicted climate. To do this on a "moving target" (as climate has a trend), I regress the future chill predictions on a third degree polynomial of years, plus a dummy variable for counties. The residuals from this regression should be free of the climate trend, and are plausibly *i.i.d.* 100 bootstraps of these residuals are added to the predicted values from the regression. The model is run, for each parametric specification, over these 100 prediction bootstraps. The expected net present values are averages over these 100 results.

A.4 FAQ about the model and solution

Figure 4.1 could raise some discomfort among economists, especially regarding the S_{MCE} curve. Below are a few explanations, presented in a Frequently Asked Questions format.

Q 1. *This S_{MCE} curve is not concave! How do you know that you will get a unique solution to the grower's problem?*

A. In short, the corner solution is $x^* = 0$ when the output price does not justify using MCE, and I can prove that there is only one profit maximizing solution for the grower. For details, please see the section: "One Optimal Solution For Grower With MCE". □

Q 2. *Why does the S_{MCE} overlap S_1 in the lower prices?*

A. In low prices, it doesn't make sense to invest in MCE because the value of marginal productivity (VMP) from it are just lower than the VMP from the ordinary input z. This is a corner solution discussed in the section: "One Optimal Solution For Grower With MCE". □

Q 3. *What happens to S_{MCE} in very high prices? It seems to get closer to S_0.*

A. Asymptotically, as the price increases S_{MCE} converges with S_0. At high prices, enough MCE has been applied so as to make the climate virtually ideal. That is, there is satiation in MCE input x. As the output price goes even higher, expenditures on MCE are virtually fixed, and further increase in production is only done by adding more input z. That is, the marginal cost becomes the same as it would be in ideal climate, a situation represented by S_0. For an extreme case, consider a grower who has MCE technologies available, but also enjoys perfect weather. In this case, S_{MCE} and S_0 are the same. □

Q 4. *Are you worried about the linear supply curves generating non-realistic grower profits in the low price range?*

A. I am definitely worried about that, but it should not matter for the grower profit **gains** from MCE technology. In the range of low prices, our modeled S_0 and S_{MCE} overlap, so all this excess modeled profit is fully deducted. In our experience, the curves start diverging at about $1,500-$2,000, a plausible marginal cost. Therefore, the grower profit gains from MCE in the model are not impacted by the linear specification in low prices. □

Appendix B

R code for simulations

B.1 Helpful functions

```
# Damage function: (actually, net of loss function)
damage_func <- function(chill, kaolin = 0){
  value <- plogis(chill + kaolin, 47.43, 8.2)
  return(value)
}

# Damage function first derivative
damage_func_dx <-  function(chill, kaolin = 0) {
  value <- dlogis(chill + kaolin, 47.43, 8.2)
  return(value)
}

# solving for supply parameters
alpha_beta_solver <- function(vec, price, quantity, elas, df){
  alpha = vec[1]
  beta2_pz = vec[2]
  df$damage = damage_func(df$chill2016)
  supply <- df$damage * df$share *
    (alpha + df$damage * beta2_pz  * price / 2)
  value1 <- quantity - sum(supply)

  dQdp <- sum(df$damage * df$share *
                ( df$damage * beta2_pz / 2))
  value2 <- dQdp * price / quantity - elas
```

```
  return(c(value1, value2))
}

# solving without kaolin (price is the solution)
no_kaolin_solver <- function(p, data_vec, nk_df){
  # data vec has some ancilliary coefficients.
  a <- data_vec[1]
  b <- data_vec[2]
  LR <- data_vec[3]
  extra_cost <- data_vec[4]

  # calculate damage for each county-quintile
  nk_df$damage <- damage_func(nk_df$chill)

  # adjust alpha and beta2 to damage
  nk_df$alpha <- nk_df$alpha * nk_df$share * nk_df$growth
  nk_df$beta2_pz <- nk_df$beta2_pz * nk_df$share * nk_df$growth

  # calculate linear supplies
  nk_df$supply <- nk_df$damage * (
    nk_df$alpha +
      nk_df$beta2_pz  * nk_df$damage * p / 2)

  # the price should zero the market number (supply - demand)
  market_num <- sum(nk_df$supply) -
    # beta "already includes" lr
    sum(nk_df$growth *  nk_df$share) * (a + b * (p + extra_cost))

  return(market_num)
}

# supply finder for kaolin treatment. To be run by optimizer
kaolin_finder <- function(kaolin_vec, data_vec, kf_df){
  # data vec has some ancilliary coefficients.
  data_vec <- as.numeric(data_vec)
  px <- data_vec[1]
  p_kaolin <- data_vec[2]

  # this should take a kf_df of one row
  # let's make it as large as the kaolin_vec
  kf_df <- kf_df %>%
   sample_n(length(kaolin_vec), replace = TRUE)
```

```
  # insert the "moving" kaolin vec (optimized on)
  kf_df$kaolin <- kaolin_vec

  kf_df$damage = damage_func(kf_df$chill, kf_df$kaolin)
  kf_df$damage_dx = damage_func_dx(kf_df$chill, kf_df$kaolin)

  # get z^* given prices
  kf_df$z_vec <- - kf_df$alpha
  kf_df$z_vec <-  kf_df$z_vec +
    sqrt(kf_df$alpha^2 +
           2 * kf_df$beta2_pz *
           (px * kf_df$acreage * kf_df$growth) *
           (1 / sqrt(kf_df$share * kf_df$growth)) *
           kf_df$damage / kf_df$damage_dx
    )
  kf_df$z_vec <- kf_df$z_vec / 2

  # get values of county necessary equations
  # in an equilibrium, these are zero
  kf_df$equation <- p_kaolin * kf_df$damage_dx *
    (kf_df$alpha * kf_df$share * kf_df$growth +
       sqrt(kf_df$share * kf_df$growth) * kf_df$z_vec) -
    px * kf_df$acreage * kf_df$growth

  return(kf_df$equation)
}


# kaolin supply builder
kaolin_supply_builder <- function(price, data_vec, ksb_df){
    # data vec has some ancillary coefficients.
  px <- data_vec[1]

  # first, try to see which county-chill-deciles can get a FOC at
     this price
  temp_list <- split(ksb_df, seq(dim(ksb_df)[1]))
  sol_list <- lapply(temp_list, function(x) try(rootSolve::uniroot
     .all(
    f = kaolin_finder,
    interval = c(-40,40),
    n = 100,
```

```
    data_vec = c(px, price),
    kf_df = x
  ),
  silent = TRUE))
  # those who get it, great. those who don't get zero kaolin
  sol_list <- lapply(sol_list, function(x) {
    if (class(x) == "try-error"){
      return(data.frame(opt_kaolin = 0))
    } else if (is_empty(x)) {
      return(data.frame(opt_kaolin = 0))
    } else if (max(x) < 0) { # this is technically impossible, but
        i keep it anyway
      return(data.frame(opt_kaolin = 0))
    } else {
      return(data.frame(opt_kaolin = max(x)))
    }
  })
  sol_list = bind_rows(sol_list)

  # optimal kaolin could be negative
  ksb_df$opt_kaolin <- as.numeric(sol_list$opt_kaolin)

  ksb_df$damage <- damage_func(ksb_df$chill, ksb_df$opt_kaolin)
  ksb_df$damage_dx <- damage_func_dx(ksb_df$chill,
     ksb_df$opt_kaolin)

  # calculating z^* for each county decile
  ksb_df$z_vec <- - ksb_df$alpha
  ksb_df$z_vec <-  ksb_df$z_vec +
    sqrt(ksb_df$alpha^2 +
           2 * ksb_df$beta2_pz *
           (px * ksb_df$acreage * ksb_df$growth) *
           (1 / sqrt(ksb_df$share * ksb_df$growth)) *
           ksb_df$damage / ksb_df$damage_dx
    )
  ksb_df$z_vec <- ksb_df$z_vec / 2

  # calculate supply (assuming positive kaolin)
  ksb_df$supply <-  ksb_df$damage * (
    ksb_df$alpha * ksb_df$share * ksb_df$growth +
      sqrt(ksb_df$share * ksb_df$growth) * ksb_df$z_vec)
```

```
  # make sure the positive kaolin solutions are local maxima
  check_vec <- exp(-6.56 + 0.141*(ksb_df$chill + ksb_df$opt_kaolin
      )) - 1
  check_vec <- check_vec / exp(-6.56 + 0.141*(ksb_df$chill +
      ksb_df$opt_kaolin))
  check_vec <- check_vec -  (ksb_df$supply - ksb_df$alpha)/
      ksb_df$supply
  stopifnot(all(check_vec[ksb_df$opt_kaolin > 0] > 0))

  # replace values of negative kaolin supplies with regular zero
      kaolin supply
  if (sum(ksb_df$opt_kaolin == 0) > 0){
    # make a vector of linear supplies
    temp_supply <- damage_func(ksb_df$chill, 0) *
      (ksb_df$alpha * ksb_df$share * ksb_df$growth +
         ksb_df$beta2_pz * ksb_df$share * ksb_df$growth *
         damage_func(ksb_df$chill, 0) * price / 2)

    ksb_df$supply[ksb_df$opt_kaolin == 0] <- temp_supply[
        ksb_df$opt_kaolin == 0]
  }

  return(ksb_df$supply)
}

# Function to solve the market with MCE.
kaolin_solver <- function(p, # the solution is the price and MCE
    levels
                          kaolin_data_vec, # more parameters
                          kaolin_df){
  # kaolin_data_vec has some ancillary coefficients.
  kaolin_data_vec <- as.numeric(kaolin_data_vec)
  a <- kaolin_data_vec[1]
  b <- kaolin_data_vec[2]
  LR <- kaolin_data_vec[3]
  px <- kaolin_data_vec[4]
  extra_cost <- kaolin_data_vec[5]

  #kaolin_supply_builder(3000, px, kaolin_df)
  # under price p, what would supply be?
  supply_vec <- kaolin_supply_builder(p, px, kaolin_df)
  # what would demand be?
```

```
  demand <- sum(kaolin_df$growth * kaolin_df$share) *
    #(a + b * LR * (p + extra_cost))
    # b already includes LR
    (a + b * (p + extra_cost))

  return(sum(supply_vec) - demand)
}

# wrapper function, returns a vector with results.
solver_wrapper <- function(data_vec, # parameter vector
                           chill_df, # full predicted chill matrix
                           share_vec,
                           acreage_vec,
                           growth_vec,
                           chill_2016, # for parameter calibration
                           price2016, quantity2016,
                           extra_cost){

  # data_vec is a row from "sim_final" data frame
  data_vec <- as.numeric(data_vec)
  ed <- data_vec[1]
  es <- data_vec[2]
  px <- data_vec[3]
  LR <- data_vec[6] # the Lerner Ratio
  year <- data_vec[7]


  ## 1) Calibrate system to get parameters
  # demand parameters
  b = ed * quantity2016 / (LR * (price2016 + extra_cost))
  a = quantity2016 - b *  LR * (price2016 + extra_cost)
  stopifnot(b < 0 & a > 0)

  ## 2) make useful data frame for chill
  chill_data <- data.frame(chill_df[chill_df$winter == year,])
  # add chill for "North" county
  chill_data <- rbind(chill_data,
                      c("North", year, rep(75, dim(chill_data)[2]
                        - 2)))
  # gather the chill data frame
  chill_data <- chill_data %>% select(-winter) %>%
    gather(key = "measure", value = "chill", -County) %>%
```

```
  arrange(County, measure)
# make sure this didn't convert chill to character
chill_data$chill <- as.numeric(chill_data$chill)
# merge the chill df with shares, acreage, and growth values
chill_data <- chill_data %>%
  left_join(acreage_vec %>%
             rename("acreage" = Acres), by= "County")

chill_data <- chill_data %>%
  left_join(share_vec %>%
             rename("share" = share_output), by = "County")

chill_data <- chill_data %>%
  left_join(growth_vec %>%
             filter(Year == year), by = "County")

chill_data <- chill_data %>%
 left_join(chill_2016, by = c("County", "measure"))

# we are doing county chill-quintiles, adjust the acreage and
   shares
chill_data$acreage <- chill_data$acreage / length(unique(
   chill_data$measure))
chill_data$share <- chill_data$share / length(unique(
   chill_data$measure))

## Get demand parameters
# initial guess
temp_beta2_pz <- 2 * quantity2016 * es / price2016
temp_alpha <- quantity2016 - 0.5 * temp_beta2_pz * price2016

# optimize to get parameters
param_solution <- rootSolve::multiroot(alpha_beta_solver,
                                       start(temp_alpha,
                                          temp_beta2_pz),
                                       price = price2016,
                                       quantity = quantity2016,
                                       elas = es,
                                       df = chill_data)

chill_data$beta2_pz <- param_solution$root[2]
chill_data$alpha <- param_solution$root[1]
```

```
stopifnot(all(chill_data$alpha > 0) & all(chill_data$beta2_pz >
   0))
rm(temp_beta2_pz, temp_alpha)

## 3) Calculate no kaolin price, quantities, CS, and profits
# first, calculate the price
p_nokaolin <- uniroot(no_kaolin_solver,
                      interval = c(0, 50000),
                     #data_vec = c(a, b, LR),
                      data_vec = c(a, b, LR, extra_cost),
                      nk_df = chill_data)$root
# calculate the supply vector w/o kaolin
nokaolin_df <- chill_data
nokaolin_df$damage <- damage_func(nokaolin_df$chill)

# adjust alpha and beta2 to shares
nokaolin_df$alpha <- nokaolin_df$alpha * nokaolin_df$share *
   nokaolin_df$growth
nokaolin_df$beta2_pz <- nokaolin_df$beta2_pz * nokaolin_df$share
    * nokaolin_df$growth

# calculate supplies
nokaolin_df$supply <- nokaolin_df$damage * (
  nokaolin_df$alpha +
    nokaolin_df$beta2_pz  * nokaolin_df$damage * p_nokaolin / 2)

# calculate county profits
# the share of "automatic" output has no MC (short run)
# this is a trapezoid, with parallel sides of supply and alpha,
   and
# height p_nokaolin

nokaolin_df$profit <- (nokaolin_df$supply + nokaolin_df$alpha *
   nokaolin_df$damage) *
  p_nokaolin / 2

# calculate CS:
# what's the choke price?
p_choke <- -a / b # that's grower price choke
# CS is the triagle area
CS_nokaolin <- sum(nokaolin_df$supply) *
```

```
  (p_choke - (LR * (p_nokaolin + extra_cost))) / 2

 #calculate intermediate profit
 inter_profit_nokaolin <- (LR - 1) * (p_nokaolin + extra_cost) *
    sum(nokaolin_df$supply)
 # welfare profits
 welfare_nokaolin <- sum(nokaolin_df$profit) + CS_nokaolin +
    inter_profit_nokaolin

 ## 4) calculate same stuff for kaolin case.

 kaolin_data  <- chill_data # just to make it different from
    previously used kaolin_df

 # solving for market price
 kaolin_solution <- uniroot(kaolin_solver, c(10, 20000),
                            kaolin_data_vec = c(a, b, LR, px,
                               extra_cost),
                            kaolin_df = chill_data)

 p_kaolin <- kaolin_solution$root
 # calculate supply
 kaolin_data$supply <- kaolin_supply_builder(p_kaolin,
                                             px, kaolin_data)

 # how much kaolin are they using at equilibrium
 # the supply builder
 temp_list <- split(kaolin_data, seq(dim(kaolin_data)[1]))
 sol_list <- lapply(temp_list, function(x)
  try(rootSolve::uniroot.all(kaolin_finder, interval = c(-50,50),
                            n = 100, data_vec = c(px, p_kaolin),
                            kf_df = x), silent = TRUE))

 # those who get it, great. those who don't get zero kaolin
 sol_list <- lapply(sol_list, function(x) {
   if (class(x) == "try-error"){
     return(data.frame(opt_kaolin = 0))
   } else if (is_empty(x)) {
     return(data.frame(opt_kaolin = 0))
   } else if (max(x) < 0) { # this is technically impossible, but
       i keep it anyway
     return(data.frame(opt_kaolin = 0))
```

```
    } else {
      return(data.frame(opt_kaolin = max(x)))
    }
  })
  sol_list = bind_rows(sol_list)


  # the optimal kaolin for those with crit_p below the price is
     zero
  kaolin_data$opt_kaolin <- as.numeric(sol_list$opt_kaolin)
  kaolin_data$actual_damage <- damage_func(kaolin_data$chill,
                                           kaolin_data$opt_kaolin)

  # calculate grower profits: approximated by rectangles on price
     vector
  price_vec <- seq(from = 0, to = p_kaolin, length.out = 20)
  price_intervals <- diff(price_vec)

  supply_df <- sapply(price_vec,
                      kaolin_supply_builder,
                      data_vec = px,
                      ksb_df = chill_data)
  supply_df <- data.frame(supply_df)
  colnames(supply_df) <- paste0("supply_", round(price_vec))
  # temp_df just helps get a vector out of it
  temp_df <- supply_df[,-1]
  temp_df <- apply(temp_df, 1, function(x, diff_vec) sum(x *
     diff_vec),
                   diff_vec = price_intervals)
  kaolin_data$profit <- temp_df

  ### calculate CS:
  # CS is the triagle area
  CS_kaolin <- sum(kaolin_data$supply) *
    (p_choke - (LR * (p_kaolin + extra_cost))) / 2

  # calculate intermediate profit
  inter_profit_kaolin <- (LR - 1) * (p_kaolin + extra_cost) * sum(
     kaolin_data$supply)

  # calculate total welfare
```

```
welfare_kaolin <- sum(kaolin_data$profit) + CS_kaolin +
  inter_profit_kaolin


##5) put it all in a vector and return.

# first, let's aggregate some stuff by county
nokaolin_df <- nokaolin_df %>% group_by(County) %>%
  summarize(
    quantity = sum(supply),
    profit = sum(profit),
    damage = mean(damage)
  ) %>% ungroup()

kaolin_data <- kaolin_data %>% group_by(County) %>%
  summarize(
    quantity = sum(supply),
    profit = sum(profit),
    damage = mean(actual_damage)
  ) %>% ungroup()

return_vec <- c(a, b,
                nokaolin_df$damage, # mean (net-of) damage rate
                # nokaolin results
                p_nokaolin,
                nokaolin_df$quantity,
                sum(nokaolin_df$quantity),
                nokaolin_df$profit,
                sum(nokaolin_df$profit),
                CS_nokaolin, inter_profit_nokaolin,
                   welfare_nokaolin,
                # kaolin results
                p_kaolin,
                kaolin_data$quantity,
                sum(kaolin_data$quantity),
                kaolin_data$profit,
                sum(kaolin_data$profit),
                CS_kaolin, inter_profit_kaolin, welfare_kaolin)

names(return_vec) <-
  c("a", "b",
    paste0(nokaolin_df$County, "_potential_damage"),
```

```
      "price_nokaolin",
      paste0(nokaolin_df$County, "_supply_nokaolin"),
      "total_supply_nokaolin",
      paste0(nokaolin_df$County, "_profit_nokaolin"),
      "profit_nokaolin",
      "CS_nokaolin", "intermediary_nokaolin", "welfare_nokaolin",
      # names for kaolin stuff
      "price_kaolin",
      paste0(kaolin_data$County, "_supply_kaolin"),
      "total_supply_kaolin",
      paste0(kaolin_data$County, "_profit_kaolin"),
      "profit_kaolin",
      "CS_kaolin", "intermediary_kaolin", "welfare_kaolin")
  print(paste0("Done with simulation: ", Sys.time()))
  return(return_vec)
}
```

B.2 Running the simulations

```
## 1) SETUP ###############################################
# clear environment
rm(list = ls())
# load packages
library(rootSolve)
library(tidyverse)

set.seed(2017)
num_sim <- 200
num_boots <- 100

## 2) PISTACHIO DATA AND GROWTH SCENARIOS ##############

# Loading pistachio yield file.
pistachio_data <- readRDS("C://users//itai.trilnick//XXXXX//data//
   pistachio_yields.rds")

## growth rates:
# I have the growth paths here:
acreage_scenarios <- pistachio_data$acreage_scenarios
# grow_high <- acreage_scenarios$High[acreage_scenarios$Year ==
   2030] /
#   acreage_scenarios$High[acreage_scenarios$Year == 2016]
```

```
# grow_low <- acreage_scenarios$Low[acreage_scenarios$Year ==
    2030] /
#   acreage_scenarios$Low[acreage_scenarios$Year == 2016]

# just making sure that the acreage share is similar to output
    share
acreage_2016 <- pistachio_data$pistachio_2016
acreage_2016$share_acres <- c(1,acreage_2016$Acres[-1] /
    acreage_2016$Acres[1])
acreage_2016 <- acreage_2016 %>% select(County, Acres, share_acres
    , share_output) %>%
  filter(County %in% c("Fresno", "Kern", "Kings", "Madera", "
      Tulare", "State Total"))
## create "north" county with rest of acreage
acreage_2016 <- data.frame(acreage_2016)
acreage_2016$Acres[acreage_2016$County == "State Total"] <-
    acreage_2016$Acres[acreage_2016$County == "State Total"] -
  sum(acreage_2016$Acres[!(acreage_2016$County == "State Total")])
acreage_2016$share_output[acreage_2016$County == "State Total"] <-
     acreage_2016$share_output[acreage_2016$County == "State Total
    "] -
  sum(acreage_2016$share_output[!(acreage_2016$County == "State
      Total")])
acreage_2016$County[acreage_2016$County == "State Total"] <- "
    North"
acreage_2016 <- acreage_2016  %>% arrange(County)

# make a df with county, year, scenario, and acreage
acres_by_year <- expand.grid(County = c("Fresno", "Kern", "Kings",
     "Madera", "Tulare", "North"), Year = 2016:2040)

acres_by_year <- left_join(acres_by_year, acreage_2016, by = "
    County")
acres_by_year <- left_join(acres_by_year, acreage_scenarios, by =
    "Year")
acres_by_year$share_output <- as.numeric(
    acres_by_year$share_output)
acres_by_year$High[is.na(acres_by_year$High)] <-
    acres_by_year$High[acres_by_year$Year == 2030]
acres_by_year$Low[is.na(acres_by_year$Low)] <- acres_by_year$Low[
    acres_by_year$Year == 2030]
acres_by_year$`No Growth`[is.na(acres_by_year$`No Growth`)] <-
```

```
   acres_by_year$`No Growth`[acres_by_year$Year == 2030]

acres_by_year$High <- acres_by_year$share_acres *
   acres_by_year$High
acres_by_year$Low <- acres_by_year$share_acres * acres_by_year$Low
acres_by_year$`No Growth` <- acres_by_year$share_acres *
   acres_by_year$`No Growth`

# now, a scenario for the "high-north"
acres_by_year$`High North` <- acres_by_year$High
acres_by_year <- acres_by_year %>%
  group_by(County) %>%
  arrange(Year) %>%
  mutate(
    High2022 = nth(High, 7)
  ) %>%
  ungroup() %>%
  mutate(
    `High North` = if_else(Year > 2022, High2022, High)
  ) %>%
  group_by(Year) %>%
  mutate(
    sum_diff = ifelse(Year > 2022, sum(High) - sum(High2022), 0),
  ) %>% ungroup()

acres_by_year$`High North`[acres_by_year$Year > 2022 &
   acres_by_year$County == "North"] <-
  acres_by_year$`High North`[acres_by_year$Year > 2022 &
     acres_by_year$County == "North"] +
  acres_by_year$sum_diff[acres_by_year$Year > 2022 &
     acres_by_year$County == "North"]

acres_by_year <- acres_by_year %>% select(County, Year,
                                          "High Same" = High,
                                          `High North`,
                                          "Low Growth" = Low,
                                          `No Growth`)
scenario_names <- c("No Growth", "Low Growth", "High North", "High
    Same")

# and now a df for shares
growth_by_year <- acres_by_year %>%
```

```
  group_by(County) %>%
  arrange(Year) %>%
  mutate(
    `High Same` = `High Same` / first(`High Same`),
    `Low Growth` = `Low Growth` / first(`Low Growth`),
    `No Growth` = `No Growth` / first(`No Growth`),
    `High North` = `High North` / first(`High North`)
  ) %>% ungroup()

## Determine price and quantity for 2016
market_data <- pistachio_data$past_data
price2016 <- market_data$Price[market_data$Year == 2016]
quantity2016 <- market_data$Production[market_data$Year == 2016]

## 3)  LOAD MAPS AND POINTS ##################################

chill_list_imported <- readRDS("XXXXX//
    final_point_data_pistachio_1km.rds")

ccsm_point_calib_chill <-
    chill_list_imported$ccsm_point_calib_chill
prob_vec <- seq(from = 10, to = 90, by = 20)/100

chill_df <- ccsm_point_calib_chill %>% select(-point_num) %>%
  group_by(County, winter) %>%
  dplyr::summarise(nest_col =
                    list(data.frame(prob = paste0("q", prob_vec),
                     chill = quantile(chill, probs = prob_vec,
                       names = FALSE)))) %>%
  unnest %>% ungroup() %>%
  spread(key = prob, value = chill)

rm(ccsm_point_calib_chill)

# get the damages in 2016
cimis_point_chill <- chill_list_imported$cimis_point_chill
chill_2016 <- cimis_point_chill %>% select(-point_num) %>%
  filter(winter == 2016) %>%
  group_by(County) %>%
  dplyr::summarise(nest_col =
                    list(data.frame(prob = paste0("q", prob_vec),
                                   chill = quantile(chill, probs
```

```
                                                = prob_vec, names = FALSE
                                               )))) %>%
  unnest %>% ungroup()
chill_2016$County <- as.character(chill_2016$County)
chill_2016$prob <- as.character(chill_2016$prob)
colnames(chill_2016) <- c("County", "measure", "chill2016")
chill_2016 <- rbind(chill_2016,
                    data.frame(
                      County = rep("North", 10),
                      measure = chill_2016$measure[1:10],
                      chill2016 = rep(75, 10)))


rm(cimis_point_chill)

## 4) SIMULATION FUNCTIONS #######

source("C://users//itai.trilnick//XXXXX//codes//
   linear_simulations_functions_NPV.R")

## 5)  Creating parameter Dataframes ######################

## main parameters: for means
sim_parameters <- expand.grid(
  ed = c(-2, -1.61, -1.1, -0.5),
  es = c(0.1, 0.19, 0.3),
  px = c(25, 55, 110),
  monopoly = c(0, 0.25, 0.5),
  monopsony = 0,
  LR = 1,
  year = 2020:2040
)
sim_parameters <- sim_parameters %>% filter(!(ed > -1 & monopoly >
    0))
# update lerner ratio
sim_parameters$LR <- 1 / (1 + (sim_parameters$monopoly /
   sim_parameters$ed))

## creating bootstrappped weather years for SE
# basically, I can't draw from the weather years because climate
   is changing over
# time. I need to de-trend the chill portions
```

```
boot_fitted <- chill_df[, 1:2]
boot_residuals <- chill_df[, 1:2]
for (col in colnames(chill_df)[-c(1,2)]){
  temp_reg <- as.formula(paste(col, "~ 0 + County + I(winter -
     2020) + I((winter - 2030)^2) + I((winter - 2030)^3)"))
  temp_reg <- lm(temp_reg, data = chill_df)
  boot_fitted$newcol <- temp_reg$fitted.values
  colnames(boot_fitted)[dim(boot_fitted)[2]] <- col
  boot_residuals$newcol <- temp_reg$residuals
  colnames(boot_residuals)[dim(boot_residuals)[2]] <- col
}


# making bootstrap weather years
boot_machine <- function(empty, df1, df2){
  #df1 is the predicted
  #df2 is residuals
  temp_df <- df2 %>% group_by(County) %>%
    sample_n(21, replace = TRUE) %>% ungroup()
  return_df <- df1
  return_df[-c(1,2)] <- return_df[-c(1,2)] + temp_df[-c(1,2)]
  return(return_df)
}
boot_list <- lapply(1:num_boots, boot_machine, boot_fitted,
   boot_residuals)

##### creating grid of outcomes for "derivative" plot
deriv_parameters <- data.frame(
                     ed = runif(num_sim, -3, -1.05),
                     es = runif(num_sim, 0.05, 0.9),
                     px = runif(num_sim, 10, 300),
                     monopoly = runif(num_sim, 0, 0.75),
                     monopsony = rep(0, num_sim),
                     LR = rep(1, num_sim)
                    )

deriv_df <- deriv_parameters
for (i in 1:20){
  deriv_df <- rbind(deriv_df, deriv_parameters)
}
deriv_df$year = rep(2020:2040, num_sim)
deriv_parameters$LR <- 1 / (1 + (deriv_parameters$monopoly /
```

```
    deriv_parameters$ed))


## 6) BASIC 2016 FIGURES FOR CALIBRATION #######

share_vec <- acreage_2016 %>% select(County, share_output)
acreage_vec <- acreage_2016 %>% select(County, Acres)
# acreage_vec <- as.numeric(acreage_df$Acres)
# counties <-
#share_vec <- c(shares_2016$share_output, 0.03)
# names(acreage_vec) <- counties
chill_df <- data.frame(chill_df)
#phi <-  0.517

# price is skewed because of what seems like a Fresno erroneous
    reporting.
price2016 <- 4300
# make price and quantity metric
price2016 <- price2016 #/ 0.91
quantity2016 <- quantity2016 #* 0.91

extra_cost <- price2016 * 0.517 # this should represent the extra
    costs in levels


####### Test runs

# data_vec <- sim_parameters[1,]
# # the starting acreage
# share_vec <- acreage_2016 %>% select(County, share_output)
# acreage_vec <- acreage_2016 %>% select(County, Acres)
# growth_vec <- growth_by_year %>% select(County, Year, growth = "
    High")
#
# temp_result <- solver_wrapper(data_vec, chill_df, share_vec,
    acreage_vec, growth_vec, chill_2016, price2016, quantity2016,
    extra_cost)
#
# print(temp_result)
#
# skipping checks (too long to include, basically make sure signs
    are ok)
```

```
rm(temp_result)


## 7) RUNNING THIS THING ##############

# real run for parameter matrix
cl <- parallel::makeCluster(8)
parallel::clusterExport(cl, c("price2016", "quantity2016", "
   growth_by_year",
                              "share_vec", "acreage_vec", "
                                 extra_cost",
                              "no_kaolin_solver", "damage_func",
                              "damage_func_dx", "
                                 kaolin_supply_builder",
                              "alpha_beta_solver",
                              "solver_wrapper", "kaolin_solver", "
                                 kaolin_finder"))
parallel::clusterEvalQ(cl, library(tidyverse))

print(paste0("Starting main results: ", Sys.time()))

for(scenario in scenario_names){
   temp_result <-  parallel::parApply(cl, sim_parameters, 1,
                        solver_wrapper,
                        chill_df,
                        share_vec = share_vec,
                        acreage_vec = acreage_vec,
                        growth_vec =  growth_by_year %>% select(
                           County, Year, growth = scenario),
                        chill_2016 = chill_2016,
                        price2016, quantity2016, extra_cost)
  temp_names <- row.names(temp_result)
  temp_result <- t(temp_result)
  temp_result <- as.data.frame(temp_result)
  colnames(temp_result) <- temp_names
  temp_result <- cbind(temp_result, sim_parameters)
  # add scenario variable
  temp_result$Scenario <- scenario
  if (scenario == "No Growth"){
    linear_result <- temp_result
  } else {
```

```
    linear_result <- rbind(linear_result, temp_result)
  }
  print(paste0("Done with ", scenario, " main results: ", Sys.time
     ()))
}
#parallel::stopCluster(cl)


##### running bootstrapped runs
# as these runs take a long time, I only bootstrap the main
   specification:
# ed = -1.61, es = 0.19, px = 55, monopoly = 0

# I have 90 parameter combinations. Now, I want to run each on a
   climate botstrap
# that is 90 X 21 (years) X 100 (bootstraps) X 4 scenarios
# each one takes about 3 seconds
# that's about 630 hours
# paralleling on 10 cores, ~63 hours
# boot seemed to run much slower. took more than 24 hours and wasn
   't finished with one scenario
# I therefore start with just the variation for elasticities


# first, a function to wrap the wrapper to get the chill data
   frame in place
boot_params <- sim_parameters %>% filter(px == 55)

boot_wrapper <- function(df, params,
                         share_vec,
                         acreage_vec,
                         growth_vec,
                         chill_2016,
                         price2016, quantity2016,
                         extra_cost) {
  # a booootstrap is a solution with 21 years. So need to run over
     all values
  # in the booted chill data frame
  temp_value <- apply(params, 1,
                      solver_wrapper,
                      chill_df = df,
                      share_vec = share_vec,
```

```
                      acreage_vec = acreage_vec,
                      growth_vec = growth_vec,
                      chill_2016,
                      price2016, quantity2016,
                      extra_cost)
  # the value are the results for all 21 years in bootstrap
  temp_names <- row.names(temp_value)
  temp_value <- t(temp_value)
  temp_value <- as.data.frame(temp_value)
  colnames(temp_value) <- temp_names
  return(temp_value)
}

parallel::clusterExport(cl, "boot_wrapper")

print(paste0("Starting main bootstraps: ", Sys.time()))

# this should return a list of data frames with simulation results
for(scenario in scenario_names){
  print(paste0("Starting bootstraps on scenario: ", scenario, "
      ", Sys.time()))
  temp_result <- parallel::parLapply(cl, boot_list,
                                      boot_wrapper,
                                      boot_params,
                                      share_vec = share_vec,
                                      acreage_vec = acreage_vec,
                                      growth_vec = growth_by_year
                                        %>% select(County, Year,
                                        growth = scenario),
                                      chill_2016 = chill_2016,
                                      price2016, quantity2016,
                                        extra_cost)
  # this returns a list
  temp_result <- bind_rows(temp_result)
  temp_params <- replicate(num_boots, boot_params,
                          simplify = FALSE)
  temp_params <- bind_rows(temp_params)
  stopifnot(dim(temp_result)[1] == dim(temp_params)[1])
  temp_result <- cbind(temp_result, temp_params)

  # add scenario variable
  temp_result$Scenario <- scenario
```

```
  if (scenario == "No Growth"){
    linear_boots <- temp_result
  } else {
    linear_boots <- rbind(linear_boots, temp_result)
  }
  print(paste0("Done with bootstraps, ", scenario, ": ", Sys.time
     ()))
}

# name the bootstrapps
boot_nums <- rep(1:num_boots, dim(boot_params)[1])
boot_nums <- boot_nums[order(boot_nums)]
#stopifnot(dim(linear_boots)[1] == length(boot_nums))
linear_boots$boot_num <- boot_nums
table(linear_boots$boot_num)
rm(boot_nums)


### now for derivatives. they are the same except for the
   parameters and the bootstraps
# I have a big matrix with varying parameters, and 21 years for
   each combination
# I'll run it on the first and second climate prediction bootstrap
print(paste0("Starting derivatives:      ", Sys.time()))
deriv_result <- parallel::parLapply(cl, boot_list[1:5],
                                    boot_wrapper,
                                    deriv_df,
                                    share_vec = share_vec,
                                    acreage_vec = acreage_vec,
                                    growth_vec = growth_by_year %>%
                                        select(County, Year, growth
                                        = `High North`),
                                    chill_2016 = chill_2016,
                                    price2016, quantity2016,
                                      extra_cost)
# this returns a list
deriv_result <- bind_rows(deriv_result)
temp_params <- replicate(5, deriv_df,
                         simplify = FALSE)
temp_params <- bind_rows(temp_params)
stopifnot(dim(deriv_result)[1] == dim(temp_params)[1])
deriv_final <- cbind(deriv_result, temp_params)
```

```
deriv_final$Scenario <- "High North"

print(paste0("Done with derivatives: ", Sys.time()))

parallel::stopCluster(cl)


# save results file
saveRDS(list("linear_main" = linear_result,
             "linear_deriv" = deriv_final,
             "linear_boots" = linear_boots,
             "chill_df" = chill_df,
             "acres_df" = acres_by_year,
             "growth_df" = growth_by_year,
             "price2016" = price2016,
             "chill_2016" = chill_2016,
             "acreage_2016" = acreage_2016,
             "quantity2016" = quantity2016,
             "README" = paste0("Linear model results, based on
                county quintile chill portions, estimated logistic
                 function", Sys.time())),
        file = "C://Users//itai.trilnick//dropbox//CIMIS project//
           itai/kaolin_paper//data//
           results_linear_bycounty_deciles_updated_allboots.rds")
```